АВТОМАТИКА
и
ТЕЛЕМЕХАНИКА

Volume 20, No. 9

September 1959

SOVIET INSTRUMENTATION AND
CONTROL TRANSLATION SERIES

Automation and Remote Control

(The Soviet Journal *Avtomatika i Telemekhanika* in English Translation)

■ This translation of a Soviet journal on automatic control is published as a service to American science and industry. It is sponsored by the Instrument Society of America under a grant in aid from the National Science Foundation, continuing a program initiated by the Massachusetts Institute of Technology.

SOVIET INSTRUMENTATION AND CONTROL TRANSLATION SERIES

■ This translation of the Soviet Journal *Avtomatika i Telemekhanika* is published and distributed at nominal subscription rates under a grant in aid to the Instrument Society of America from the National Science Foundation. This translated journal, and others in the Series (see back cover), will enable American scientists and engineers to be informed of work in the fields of instrumentation, measurement techniques, and automatic control reported in the Soviet Union.

The original Russian articles are translated by competent technical personnel. The translations are on a cover-to-cover basis, permitting readers to appraise for themselves the scope, status, and importance of the Soviet work.

Publication of *Avtomatika i Telemekhanika* in English translation started under the present auspices in April, 1958, with Russian Vol. 18, No. 1 of January, 1957. Translations of Vols. 18 and 19 have now been completed. Translation of Vol. 20, No. 1 will be published in January, 1960, and the twelve issues of Vol. 20 will be published in English translation by October, 1960.

Transliteration of the names of Russian authors follows the system known as the British Standard. This system has recently achieved wide adoption in the United Kingdom, and is being adopted in 1959 by a large number of scientific journals in the United States.

All views expressed in the translated material are intended to be those of the original authors, and not those of the translators, nor the Instrument Society of America.

Readers are invited to submit communications on the quality of the translations and the content of the articles to ISA headquarters. Pertinent correspondence will be published in the "Letters" section of the ISA Journal. Space will also be made available in the ISA Journal for such replies as may be received from Russian authors to comments or questions by American readers.

1959 Volume 20 Subscription Prices:

Per year (12 issues), starting with Vol. 20, No. 1

General:	*United States and Canada*	$35.00
	Elsewhere	38.00
Libraries of nonprofit academic institutions:		
	United States and Canada	$17.50
	Elsewhere	20.50
Single issues to everyone, each		$ 6.00

1957 Volume 18 and 1958 Volume 19 issues also available. Prices upon request.

See back cover for combined subscription to entire Series.

Subscriptions and requests for information on back issues should be addressed to the:

Instrument Society of America
313 Sixth Avenue, Pittsburgh 22, Penna.

Translated and printed by Consultants Bureau Enterprises, Inc.

Volume XX No. 9 September 1959

English Translation Published May 1960

Automation and Remote Control

The Soviet Journal Avtomatika i Telemekhanika in English Translation

Reported circulation of the Russian original 8,000.

Avtomatika i Telemekhanika is a Publication of the Academy of Sciences of the USSR

See following page for Table of Contents.

AUTOMATION AND REMOTE CONTROL

Volume 20, Number 9 **September, 1959**

CONTENTS

ON THE STABILITY OF A RELAY SYSTEM'S EQUILIBRIUM STATE

S. D. Kinyapin and Yu. I. Neimark
(Gor'kii)

A number of works have been devoted to the investigation of the stability of a relay system's equilibrium state. Initially, stability conditions obtained on the basis of inductive reasoning and the results of M. V. Meerov's investigations were cited by Ya. Z. Tsypkin [1]. Subsequently, they received a basis in the work of L. S. Pontryagin and V. G. Boltyanskii, a brief discussion of which is contained in [2]* and, in the case when the difference of the orders of the denominator and numerator of the relay system's linear link's transfer constant is two, in work [3].

In the present paper a new derivation of the stability conditions for a relay system's equilibrium state is given. With this, the approach to the question is borrowed from work [3] but the investigation of the transformation's fixed points' stability is carried out in a different way, being based on the relationship, established in [4], between the stability of the equilibrium state of the system of differential equations and the stability of the fixed points of the corresponding mapping.

Preliminary Remarks and Posing of the Problem; Information About the Point Mapping

If the transfer constant, K(p), or the admittance function, $\varphi(t)$, of a relay system's linear link, shown on Fig. 1 with the relay characteristic shown in Fig. 2, satisfy the conditions

$$\lim_{p\to\infty} K(p) = \varphi(0) = 0, \quad \lim_{p\to\infty} pK(p) = \varphi'(0) < 0, \tag{1}$$

then, as is well known [5], the so-called sliding motion is possible in the relay system.

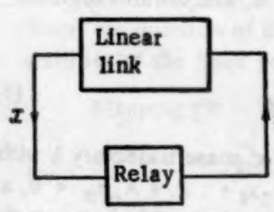


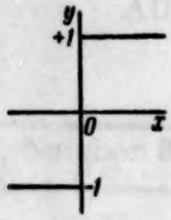
Fig. 1.

If conditions (1) hold, then, on the relay switching plane in the relay system's phase space, there is a so-called sliding motion lamina. The equilibrium state lies on the sliding motion lamina, and is stable if all the zeroes of K(p) lie in the left hand-plane, and unstable otherwise [5]. The phase trajectories close to, and on, the lamina are represented in Fig. 3.

As $\varphi'(0)$ goes to zero, the sliding motion lamina disappears, and sliding mode operation can no longer occur. Convergence of the phase points to the equilibrium state is only possible for ever accelerating relay switching, as shown on Fig. 4. More exactly, in order that the equilibrium state be asymptotically stable, it is necessary that, first of all, for the phase points which are approaching the equilibrium position the time interval between successive relay switchings decrease without bound and, secondly, the points of intersection of the phase trajectory with the relay switching plane, M_0, M_1, M_2, . . . , converge to the equilibrium position for those points which are in the neighborhood of the equilibrium position.

*After this paper was sent to the printer, work [8] was published, introducing simplifications in the investigations of L. S. Pontryagin and V. G. Boltyanskii.

We assume that



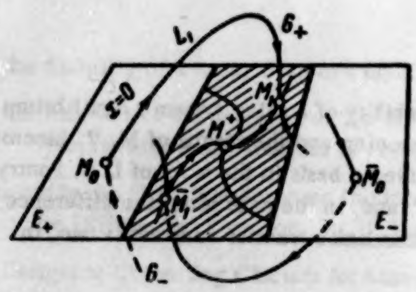
$$K(p) = \frac{A_m(p)}{B_n(p)} = \sum_{i=1}^{n} \frac{c_i}{p + \lambda_i}, \tag{2}$$

where

Fig. 2.

$$A_m(p) = a_0 p^m + a_1 p^{m-1} + \ldots + a_m, \quad a_0 \neq 0, \tag{3}$$
$$B_n(p) = p^n + b_1 p^{n-1} + \ldots + b_n,$$

and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are roots of the polynomial $B_n(p)$. Since we are interested in the case when $\varphi(0) = \varphi'(0)$, then $m \leq n - 2$ and



$$c_1 + c_2 + \ldots + c_n = 0. \tag{4}$$

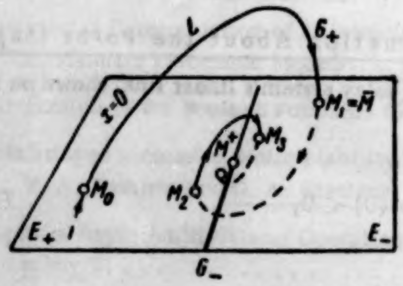
The equation of motion of the relay system under consideration may be written in the form* [5, 6]

$$\dot{x}_i + \lambda_i x_i = c_i \operatorname{sign}(x_1 + x_2 + \ldots + x_n). \tag{5}$$

Fig. 3.

We take $x_1, x_2, \ldots, x_n$ as the phase variables. In the region G_+, where $x_1 + x_2 + \ldots + x_n > 0$, and in region G_-, where $x_1 + x_2 + \ldots + x_n < 0$, the solution of equations (5) has the form:

$$x_i = x_i^0 e^{-\lambda_i (t - t_0)} \pm \frac{c_i}{\lambda_i} (1 - e^{-\lambda_i (t - t_0)}), \tag{6}$$



where $x_i^0 = x_i(t_0)$ and the upper sign should be chosen for region G_+ the lower sign for region G_-.

Let $M(x_1, x_2, \ldots, x_n)$ be a point lying on the relay switching plane

$$x = x_1 + x_2 + \ldots + x_n = 0. \tag{7}$$

Fig. 4.

A phase trajectory passing through point M is directed towards region G_+ if $\dot{x} > 0$, and towards region G_- if $\dot{x} < 0$. By adding equations (5) termwise, and keeping (4) in mind, we find that

$$\dot{x} = \dot{x}_1 + \ldots + \dot{x}_n = -\lambda_1 x_1 - \lambda_2 x_2 - \ldots - \lambda_n x_n. \tag{8}$$

Let $M(x_1, \ldots, x_n)$ and $\overline{M}(\overline{x}_1, \ldots, \overline{x}_n)$ be two successive points of intersection of phase trajectory L with the relay switching plane $x = 0$. Segment $M\overline{M}$ of the phase trajectory will lie in G_+ if $\lambda_1 x_1 + \ldots + \lambda_n x_n < 0$, and in G_- for the inverse inequality. In accordance with (6), there will hold the following relationship between the coordinates of points M and $\overline{M}$

$$\overline{x}_i = x_i e^{-\lambda_i \tau} \pm \frac{c_i}{\lambda_i} (1 - e^{-\lambda_i \tau}), \tag{9}$$

---

*By definition, $\operatorname{sign} \xi = \begin{cases} +1 \text{ for } \xi > 0, \\ -1 \text{ for } \xi < 0. \end{cases}$

where τ, the time of motion of the phase point from M to $\overline{M}$, is the least positive root of the equation

$$\sum x_i e^{-\lambda_i \tau} \pm \frac{c_i}{\lambda_i}(1 - e^{-\lambda_i \tau}) = 0. \tag{10}$$

In (9) and (10), the upper or lower sign is chosen as a function of the sign of the quantity $\lambda_1 x_1 + \ldots + \lambda_n x_n$. Equation (10) can be written in the form

$$\varphi(\tau) = \theta(\tau),$$

where

$$\theta(\tau) = \mp \sum x_i e^{-\lambda_i \tau} \text{and} \left(\frac{d\theta}{d\tau}\right)_{\tau=0} = \mp \sum -\lambda_i x_i < 0.$$

In order that, in a neighborhood close to the equilibrium position, $x_1 = \ldots = x_n = 0$, equation (10) have positive roots which tend to zero as point M approaches the equilibrium point, the function $\varphi(t)$ must be negative for sufficiently small $\underline{t}$. Since, for small $\underline{t}$

$$\varphi(t) = -\frac{d_1}{2} t - \frac{d_2}{6} t^2 - \frac{d_3}{24} t^3 - \ldots,$$

where

$$d_s = (-1)^{s-1} \sum c_i \lambda_i^s, \tag{11}$$

the first nonzero coefficient d_k must be positive, i.e.,

$$\text{if } d_1 = d_2 = \ldots = d_{k-1} = 0, \text{then } d_k > 0. \tag{12}$$

Relationships (12) provide a necessary condition for the stability of the equilibrium state in the relay's switching mode. If conditions (12) hold, then, for stability in the small of the equilibrium state M, it is necessary and sufficient that the sequence of points, M, $\overline{M}$, $\overline{\overline{M}}$, ..., defined by relationship (9), tend to equilibrium state M^* when point M is sufficiently close to M^*.

Relationship (9) may be considered as a point mapping of the switching plane into itself, translating point M to $\overline{M}$. The point M^*, with coordinates $x_1 = x_2 = \ldots = x_n = 0$, is mapped into itself by this transformation. Thus, the question of the stability of equilibrium point M^* of the relay system reduces to the question of the stability of the fixed point M^* of mapping (9).

Mapping (9) maps every point for which one of the following inequalities holds

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n < 0, \; \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n > 0, \tag{13}$$

into a point for which the inverse inequality holds. Therefore, if relay switching plane x = 0 is subdivided into two half-planes, E_+ and E_-, in which the first or, respectively, the second of the inequalities in (13) holds, then mapping (9) will map half-plane E_+ on E_-, and half-plane E_- on E_+. If we now introduce the new coordinates $X_i' = -x_i$ in half-plane E_-, then the transformations of E_+ into E_- and of E_- into E_+ will have the identical form. Specifically,

$$x_i' = -x_i e^{-\lambda_i \tau_1} - \frac{c_i}{\lambda_i}(1 - e^{-\lambda_i \tau}) \text{and } x_i = -x_i' e^{-\lambda_i \tau_2} - \frac{c_i}{\lambda_i}(1 - e^{-\lambda_i \tau_2}),$$

where τ_1 and τ_2 are the least positive roots of the equations, respectively, $\Sigma x_i' = 0$ and $\Sigma x_i = 0$.

Therefore, the question of the stability of the fixed point M^* of transformation (9) reduces to the investigation of stability of the fixed point of the transformation

$$\bar{x}_i = -x_i e^{-\lambda_i \tau} - \frac{c_i}{\lambda_i}(1 - e^{-\lambda_i \tau}), \tag{14}$$

where τ is the least positive root of the equation

$$\sum x_i e^{-\lambda_i \tau} + \frac{c_i}{\lambda_i}(1 - e^{-\lambda_i \tau}) = 0), \tag{15}$$

where, for the variables x_i, the following inequality holds

$$\sum \lambda_i x_i < 0. \tag{16}$$

Investigation of the Stability of the Fixed Point

By expanding the right members of (14) and equation (15) in series in terms of τ, we find that

$$\bar{x}_i = -x_i(1 - \lambda_i \tau + \ldots) - c_i \tau \left(1 - \frac{\lambda_i \tau}{2} + \ldots\right), \tag{17}$$

$$\sum x_i(1 - \lambda_i \tau + \ldots) + c_i \tau \left(1 - \frac{\lambda_i \tau}{2} + \ldots\right) = 0. \tag{18}$$

For the sequel, it is convenient to go to the new variables

$$u_j = \sum x_i \lambda_i^j. \tag{19}$$

If we multiply the first equation of (17) by λ_1^j, the second by λ_2^j, etc. and then add termwise, we find that

$$\bar{u}_j = -u_j + (-1)^{j-1} \tau d_j + \tau u_{j+1} + (-1)^j \frac{\tau^2}{2} d_{j+1} + \ldots, \tag{20}$$

where the repeated dots indicate terms of higher than second order in $u_1, u_2, \ldots, u_{n-1}$ and τ. In the new variables, the equation for τ takes the form:

$$-u_1 - \frac{\tau}{2} d_1 + \frac{\tau}{2} u_2 + \ldots = 0. \tag{21}$$

Further investigation of the stability of fixed point M^*, with the coordinates $u_1 = u_2 = \ldots u_{n-1} = 0$, will be carried out separately for the cases $n-m = 2$ and $n-m \geq 3$.

1. For $n-m = 2$ and $d_1 > 0$, we find from (21) that

$$\tau = -\frac{2u_1}{d_1} - \frac{4}{3}\frac{d_2}{d_1^3} u_1^2 - 2\frac{u_1 u_2}{d_1^2} + \ldots \tag{22}$$

After substitution of (22) in (20), we go to the following formulae for the transformation T under consideration by us:

$$\bar{u}_1 = u_1 - \frac{2}{3}\frac{d_2}{d_1^2}u_1^2 + \ldots,$$

$$\bar{u}_2 = -u_2 - 2\frac{d_2}{d_1}u_1 - 2\frac{d_2}{d_1^2}u_1 u_2 - 2\frac{u_1 u_3}{d_1} + 2\frac{u_1^2}{d_1^2}\left(d_3 - \frac{2}{3}\frac{d_2^2}{d_1}\right) + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{u}_{n-1} = -u_{n-1} + (-1)^n\frac{2d_{n-1}}{d_1}u_1 + (-1)^n\frac{2d_{n-1}}{d_1^2}u_1 u_2 - 2\frac{u_1 u_n}{d_1} + \tag{23}$$

$$+ (-1)^{n-1}\frac{2u_1^2}{d_1^2}\left(d_n - \frac{2}{3}\frac{d_2 d_{n-1}}{d_1}\right) + \ldots$$

It is easily seen that the general method [4] of investigating the stability of the fixed point M of transformation (23) does not lead to the goal, since all roots of the characteristic equation have unit modulus. In this particular case, investigation of stability can be carried out in the following way. We set up the transformation T^2:

$$\bar{\bar{u}}_1 = \bar{u}_1 - \frac{2}{3}\frac{d_2}{d_1^2}\bar{u}_1^2 = u_1 - \frac{4}{3}\frac{d_2}{d_1^2}u_1^2 + \ldots,$$

$$\bar{\bar{u}}_2 = -\bar{u}_2 + \ldots = u_2 + 4\frac{d_2}{d_1^2}u_1 u_2 + 4\frac{u_1 u_3}{d_1} + \frac{4}{3}\frac{d_2^2 u_1^2}{d_1^3} + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{24}$$

$$\bar{\bar{u}}_{n-1} = -\bar{u}_{n-1} + \ldots = u_{n-1} + (-1)^n\frac{4d_{n-1}}{d_1^2}u_1 u_2 + 4\frac{u_1 u_n}{d_1} +$$

$$+ (-1)^{n-1}\frac{4d_2 d_{n-1}}{3d_1^3}u_1^2 + \ldots$$

In accordance with § 8 of work [4], the fixed point of transformation T^2 (and, consequently, of T also) is stable or unstable depending on whether the equilibrium state M^* is stable or unstable in the linear approximation of the system of differential equations

$$\frac{du_1}{dt} = -\frac{4}{3}\frac{d_2}{d_1^2}u_1\frac{u_1}{\tau} + \ldots,$$

$$\frac{du_2}{dt} = 4\frac{d_2}{d_1^2}u_2\frac{u_1}{\tau} + 4\frac{u_3}{d_1}\frac{u_1}{\tau} + \frac{4d_2^2 u_1}{3d_1^3}\frac{u_1}{\tau} + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{25}$$

$$\frac{du_{n-1}}{dt} = (-1)\frac{4d_{n-1}u_2}{d_1^2}\frac{u_1}{\tau} + 4\frac{u_n}{d_1}\frac{u_1}{\tau} + (-1)^{n-1}\frac{4d_2 d_{n-1}u_1}{3d_1^3}\frac{u_1}{\tau} + \ldots$$

obtained from formula (24) for transformation T^2 by replacing $\frac{\bar{\bar{u}}_i - u_i}{\tau}$ by $\frac{du_i}{dt}$. The stability of the equilibrium state of system (25) can be investigated by the ordinary means. Specifically, by noting that, according to (22)

$$\frac{u_1}{\tau} = -\frac{d_1}{2}\left(1 - \frac{2}{3}\frac{d_2}{d_1^2}u_1 - \frac{u_2}{d_1} + \ldots\right) \tag{26}$$

and that, moreover, the following relationship holds (cf. Appendix 1) between the variables $u_1, u_2, \ldots u_n$

$$u_n - b_1 u_{n-1} + \ldots + (-1)^{n-1} b_{n-1} u_1 + (-1)^n b_n u_0 = 0, \tag{27}$$

we write the characteristic equation of system (25) of differential equations in the form

$$\Delta(p) = \begin{vmatrix} -\frac{1}{3}\frac{d_2}{d_1} + p & 0 & 0\;0 \ldots & 0 & 0 \\ \frac{1}{3}\frac{d_2^2}{d_1^2} & \frac{d_2}{d_1} + d & 1\;0 \ldots & 0 & 0 \\ -\frac{1}{3}\frac{d_2 d_3}{d^2} & -\frac{d_3}{d_1} & p\;1 \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (-1)^{n-1}\frac{d_2 d_{n-1}}{3d_1^2} & (-1)^n \frac{d_{n-1}}{d_1} & 0\;0 \ldots & p & 1 \\ (-1)^{n-1} b_{n-1} & (-1)^{n-2} b_{n-2} & \ldots & -b_1 & 1 \end{vmatrix} \tag{28}$$

By expanding determinant (28) (cf. Appendix 2), we find that

$$\Delta(p) = \left(-\frac{1}{3}\frac{d_2}{d_1} + p\right) A_m(p), \tag{29}$$

where $A_m(p)$ is the numerator of the transfer constant of the relay system's linear link. In accordance with what has already been said, and in correspondence with (29), the equilibrium state for the relay system for $n - m = 2$ will be stable if

$$d_1 > 0, \; -d_2 > 0 \tag{30}$$

and if, in addition, all the roots of $A_m(p)$ lie in the left half-plane. By expressing d_1 and d_2 in terms of the coefficients of polynomials $A_m(p)$ and $B_n(p)$, we find that (30) is equivalent to the conditions

$$a_0 < 0, \; a_1 - a_0 b_1 > 0. \tag{31}$$

2. We now turn to the consideration of the case when $n - m \geq 3$. For $n - m = 3$, $d_1 = 0$ but $d_2 \neq 0$. We introduce the new variable $v_1 = u_1/\tau$ and write equations (20) and (21) in the form

$$\begin{aligned} \bar{v}_1 &= -v_1 + u_2 - \frac{\tau}{2} d_2 - \frac{\tau}{2} u_3 - \frac{\tau^2}{6} d_3 + \tau^2(\ldots), \\ \bar{u}_2 &= -u_2 + \tau d_2 + \tau u_3 + \frac{\tau^2}{2} d_3 + \tau^2(\ldots), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \bar{u}_{(n-1)} &= -u_{n-1} + (-1)^n \tau d_{n-1} + \tau u_n + (-1)^{n-1}\frac{\tau^2}{2} d_n + \tau^2(\ldots), \end{aligned} \tag{32}$$

$$-v_1 + \frac{u_2}{2} - \frac{\tau}{6} d_2 - \tau(\ldots) = 0. \tag{33}$$

We find from equations (33) that

$$\tau = -\frac{6}{d_2}\left(v_1 - \frac{u_2}{2}\right) + \ldots$$

By substituting the value of τ just found in (32), we are led to the following formulae for the transformation T being considered by us:

$$\bar{v}_1 = 2v_1 - \frac{1}{2}u_2 + \ldots,$$

$$\bar{u}_2 = -6v_1 + u_2 + \ldots, \tag{34}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{u}_{n-1} = -u_{n-1} + (-1)^n \frac{6d_{n-1}}{d_2}\left(v_1 - \frac{u_2}{2}\right) + \ldots$$

The characteristic equation for transformation (34) is written in the form

$$X(z) = \left|\begin{array}{cc|cccc} 2-z, & -\frac{1}{2} & 0, & 0, & \ldots & 0 \\ -6, & 2-z & 0, & 0, & \ldots & 0 \\ \hline \ldots & \ldots & -1-z, & 0 & \ldots & 0 \\ \ldots & \ldots & 0, & -1-z, & \ldots, & 0 \\ & & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & 0, & 0, & \ldots, & -1-z \end{array}\right| \tag{35}$$

It has the (n −3)-fold root −1 and, in addition, the two roots defined by the equation

$$z^2 - 4z + 1 = 0.$$

One of the roots of this quadratic equation is greater than unity in modulus and, consequently, the fixed point $v_1 = u_2 = \ldots = u_{n-1} = 0$ of transformation (34) is unstable. From the instability of transformation (34) (since, with this, not only v_1 but also u_2 increases) flows the instability of the fixed point of transformation T and, consequently, the instability as well as of the equilibrium state of the relay system being considered.

In the case when $n - m = k$, for any $k \geq 3$, we proceed analogously. Specifically, as a result of the replacements $v_1 = \frac{u_1}{\tau^{k-1}}$, $v_2 = \frac{u_2}{\tau^{k-2}}, \ldots, v_{k-1} = \frac{u_{k-1}}{\tau}$, and considering that $d_1 = d_2 = \ldots = d_{k-1} = 0$ and $d_k \neq 0$, we write formulae (20) and (21) in the form

$$\bar{v}_1 = -v_1 + v_2 - \frac{v_3}{2} + \ldots + \frac{(-1)^k u_k}{(k-1)!} - \frac{\tau d_k}{k!} + \ldots,$$

$$\bar{v}_2 = -v_2 + v_3 - \frac{v_4}{2} + \ldots + \frac{(-1)^{(k-1)} u_k}{(k-2)!} + \frac{\tau d_k}{(k-1)!} \ldots, \tag{36}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{v}_{k-1} = -v_{k-1} + v_k + (-1)^{k-1}\frac{\tau d_k}{2} + \ldots,$$

$$\bar{u}_k = -u_k + (-1)^k \tau d_k + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{u}_{n-1} = -u_{n-1} + (-1)^{n-1}\tau d_{n-1} + \ldots,$$

$$v_1 - \frac{v_2}{2} + \frac{v_3}{3!} + \ldots + \frac{(-1)^{k-1}u_k}{k!} + \frac{\tau d_k}{(k+1)!} + \ldots = 0. \tag{37}$$

By substituting in (36) the quantity

$$\tau = -\frac{(k+1)!}{d_k}\left(v_1 - \frac{v_2}{2} + \frac{v_3}{3!} + \ldots + \frac{(-1)^{k-1}u_u}{k!}\right) + \ldots,$$

defined by (37), we write transformation T in the new variables in the form

$$\bar{v}_1 = kv_1 - \frac{(k-1)\,v_2}{2} + \ldots + \frac{(-1)^{i+1}\,v_i\,(k+1-i)}{i!} + \ldots + \frac{(-1)^{k+1}u_k}{k!} + \ldots,$$

$$\bar{v}_2 = -(k+1)\,kv_1 + \ldots + \frac{(-1)^{i+2}\,v_i}{i!}\,[(k+1)\,k - i\,(i-1)] +$$
$$+ \ldots + \frac{(-1)^{k+2}\,2u_k}{(k-1)!} + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{v}_{k-1} = (-1)^k\,\frac{(k+1)!}{2}\,v_1 + \ldots + \frac{(-1)^{k+i-1}\,v_i}{i!}\,[(k+1)\,k\ldots 3 - i\,(i-1)\ldots \tag{38}$$
$$\ldots(i-(k-2))] + \ldots - \frac{(k-1)\,u_k}{2} + \ldots,$$

$$\bar{u}_k = (-1)^{k+1}\,(k+1)!\,v_i + \ldots + \frac{(-1)^{k+i}\,v_i}{i!}\,[(k+1)! - i\,(i-1)\ldots$$
$$\ldots(i-k+1)] + \ldots + ku_k + \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\bar{u}_{n-1} = -u_{n-1} - (-1)^k\,(k+1)!\left(v_1 - \frac{v_2}{2} + \ldots + \frac{(-1)^{k-1}\,u_k}{k!}\right) + \ldots$$

The characteristic equation of transformation T,

$$X(z) = \left|\begin{array}{cccc|ccc} k-z & -\dfrac{k-1}{2} & \ldots & \dfrac{(-1)^{k+1}}{k!} & 0 & \ldots & 0 \\ -(k+1)\,k & \dfrac{(k+1)\,k-2}{2} - z & \ldots & \dfrac{(-1)^{k+2}.2}{(k-1)!} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ (-1)^{k+1}\,(k+1)! & \dfrac{(-1)^{k+2}\,(k+1)!}{2} & \ldots & k-z & 0 & \ldots & 0 \\ \hline \ldots & \ldots & \ldots & \ldots & -1-z, & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0, & \ldots & -1-z \end{array}\right| \tag{39}$$

has the (n −k + 1)-fold root −1, with the remaining m roots defined by the equation

$$X^k(z) = \left|\begin{array}{cccc} k-z & -\dfrac{k-1}{2} & \ldots\dfrac{(-1)^k\cdot 2}{(k-1)!} & \dfrac{(-1)^{k+1}}{k!} \\ -(k+1)\,k & \dfrac{(k+1)\,k-2}{2} - z & \ldots\ldots\ldots & \dfrac{(-1)^{k+2}.2}{(k-1)!} \\ \ldots & \ldots & \ldots & \ldots \\ (-1)^{n+1}\,(k+1)! & \dfrac{(-1)^{k+2}\,(k+1)!}{2} & \ldots -(k+1)\,k & k-z \end{array}\right| \tag{40}$$

which has a root greater than unity in absolute value, since the coefficient of z^{k-1} is greater than k. Thus, for any $k \geq 3$, there is instability.

Appendix 1

From the joint holding of the equations

$$
\begin{aligned}
&x_1 + x_2 + \ldots + x_n + x_{n+1}u_0 = 0,\\
&x_1\lambda_1 + x_2\lambda_2 + \ldots + x_n\lambda_n + x_{n+1}u_1 = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&x_1\lambda_1^n + x_2\lambda_2^n + \ldots + x_n\lambda_n^n + x_{n+1}u_n = 0
\end{aligned}
$$

it follows that

$$
\Delta_0 = \begin{vmatrix} 1 & 1 & \ldots & 1 & u_0 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n & u_1 \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 & u_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \lambda_1^n & \lambda_2^n & \ldots & \lambda_n^n & u_n \end{vmatrix} = 0.
$$

By expanding Δ_0 by the last column, and by introducing the notation

$$
\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix} = \Delta_{n-1} \quad \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^{s-1} & \lambda_2^{s-1} & \ldots & \lambda_n^{s-1} \\ \lambda_1^{s+1} & \lambda_2^{s+1} & \ldots & \lambda_n^{s+1} \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^n & \lambda_2^n & \ldots & \lambda_n^n \end{vmatrix} = \Delta_n^s,
$$

we have

$$
\Delta_0 = u_0\Delta_n^1 - u_1\Delta_n^2 + \ldots + (-1)^n u_n \Delta_n^{n+1} = 0.
$$

We note that $\Delta_n^s = \sigma_{n-s}\Delta_{n-1}$ and that σ_{n-s} is the sum of all possible products of $\underline{s}$ elements of the $\underline{n}$ quantities $\lambda_1, \lambda_2, \ldots, \lambda_n$. Therefore, by bearing in mind that σ_{n-s} is the coefficient b_{n-s+1} of polynomial $B_n(p)$, we finally obtain

$$
u_0 b_n + (-1)^1 u_1 b_{n-1} + \ldots + (-1)^n u_n = 0.
$$

Appendix 2

$$
\Delta(p) = \left(p - \frac{1}{3}\frac{d_2}{d_1}\right)\Delta^1(p).
$$

where

$$\Delta^1(p) = \begin{vmatrix} (-1)^{n-2}b_{n-2} & (-1)^{n-3}b_{n-3} & (-1)^{n-4}b_{n-4} & \dots & -b_1 & 1 \\ p+\frac{d_2}{d_1} & 1 & 0 & \dots & 0 & 0 \\ -\frac{d_3}{d_1} & p & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ (-1)^n\frac{d_{n-1}}{d_1} & 0 & 0 & \dots & p & 1 \end{vmatrix} \tag{I}$$

By expanding the determinant $\Delta^1(p)$ by the first column, we find that it equals

$$(-1)^{n-2}b_{n-2} - \left(p+\frac{d_2}{d_1}\right)\begin{vmatrix} (-1)^{n-3}b_{n-3} & (-1)^{n-4}b_{n-4} & \dots & -b_1 & 1 \\ p & 1 & \dots & 0 & 0 \\ 0 & p & 1\dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p & 1 \end{vmatrix} -$$

$$-\frac{d_3}{d_1}\begin{vmatrix} (-1)^{n-3}b_{n-3} & (-1)^{n-4}b_{n-} & \dots & -b_1 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & p & 1\dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p & 1 \end{vmatrix} - \dots \tag{II}$$

$$\dots - \frac{d_{n-1}}{d_1}\begin{vmatrix} (-1)^{n-3}b_{n-3} & (-1)^{n-4}b_{n-4} & \dots & -b_1 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ p & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p\,1 & 0 \end{vmatrix} =$$

$$= (-1)^{n-2}b_{n-2} - \left(p+\frac{d_2}{d_1}\right)\Delta_3 - \frac{d_3}{d_1}\Delta_4 - \dots - \frac{d_{n-1}}{d_1}\Delta_n,$$

where

$$\Delta_k = \begin{vmatrix} (-1)^{n-k}b_{n-k} & (-1)^{n-k-1}b_{n-k-1} & \dots & -b_1 & 1 \\ p & 1 & \dots & 0 & 0 \\ 0 & p & 1\dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p & 1 \end{vmatrix}$$

In accordance with [7]

$$\Delta_k = (-1)^{n-k}\,[b_{n-k} + b_{n-k-1}p + b_{n-k-2}p^2 + \dots + b_1p^{n-k-1} + p^{n-k}]. \tag{III}$$

After substitution of (III) in (II) and dropping of the nonessential factor $(-1)^{n-2}d_1^{-1}$, we get

$$\begin{aligned} \Delta e = d_1p^{n-2} + \dots + p^k\ [d_1b_{n-k-2} + d_2b_{n-k-3} + \dots + d_{n-k-1}] + \dots \\ \dots + p\,[d_1b_{n-3} + d_2b_{n-4} + \dots + d_{n-2}] + d_1b_{n-2} + b_{n-3}d_2 + \dots + d_{n-1}. \end{aligned} \tag{IV}$$

By expressing the coefficients b_i and d_i in terms of $c_1, c_2, \dots, c_n$ and $\lambda_1, \lambda_2, \dots, \lambda_n$ and by using expression (4) in expressing the c_i, we get that polynomial $\Delta^1(p)$ coincides with polynomial $A_m(p)$, written in the form

$$A_m(p) = c_1(p+\lambda_2)(p+\lambda_3)\ldots(p+\lambda_n) + (p+\lambda_1)c_2(p+\lambda_3)\ldots(p+\lambda_n) + \\ + \ldots\ldots\ldots + (p+\lambda_1)(p+\lambda_2)\ldots(p+\lambda_{n-1})c_n.$$

Received June 2, 1958

LITERATURE CITED

[1] Ya. Z. Tsypkin, Theory of Relay Automatic Control Systems [in Russian] (Moscow, 1955).

[2] R. V. Gamkrelidze Seminar of L. S. Pontryagin on Mathematical Problems in the Theory of Oscillations and Automatic Control [in Russian] UMN, 12, 3 (1957).

[3] Yu. I. Neimark, "The dynamics of relay automatic control systems," [in Russian], Dissertation (1956).

[4] Yu. I. Neimark, "The method of point transformations in the theory of nonlinear oscillations," [in Russian] Izb. Vysshei shkoly, radiofizika, No. 1 (1958).

[5] Yu. I. Neimark, "On the sliding mode of relay automatic control systems," [in Russian] Automation and Remote Control (USSR) 18, 1 (1957). "On the sliding mode and periodic motions of relay systems," [in Russian]. Trudy GIFTI i Radiofiz. Inst. GGU. Uchenye zapiski, volume XXX, physical series (1956).

[6] A. I. Lur'e, Certain Nonlinear Problems in Automatic Control Theory [in Russian] (Moscow, Leningrad, 1951).

[7] D. K. Faddeev and I. S. Sominskii, Collected Problems in Higher Algebra [in Russian].

[8] D. V. Anosov, "On the stability of relay systems' equilibrium positions," [in Russian]. Automation and Remote Control (USSR) 20, 2 (1959).

ON THE ANALYSIS OF THE STABILITY OF THE PERIODIC MODES OF OPERATION IN NONLINEAR CONTROL SYSTEMS WITH MANY DEGREES OF FREEDOM

V. A. Taft

(Moscow)

On the basis of a generalized corollary of Hill's equation, a derivation is given of the characteristic equation (in finite form) of a system with many degrees of freedom whose parameters are periodic functions of time. The results obtained permit the use of the well-known Mikhailov criterion for analyzing the stability of the periodic modes of operation.

The question as to the stability of periodic modes of operation is quite essential for various technical problems. For example, in electrical engineering, an analysis of the existence of periodic modes turns out to be necessary for the investigation of the modes of operation of assemblies transmitting dc and ac energy (the stability of operation of rectifying and inverting units, the generation of subharmonic oscillations, etc.), for the investigation of the parametric generation of ac currents, and elsewhere. In automatic control theory, the investigation of periodic modes is important for the analysis of operation of vibration regulators with account taken of nonlinearities and other factors, and in radio engineering for the analysis of oscillation generator operation.

The investigation of the stability of periodic modes of operation leads to the investigation of the stability of solutions of differential equations with periodic coefficients. A large number of works are devoted to this question, but there is no exact solution for the general case. Relatively recently there were made attempts to apply, to the solution of this problem, well-known graphical methods based on frequency methods [1, 2].

In particular, the problem of obtaining an approximate solution based on the assumption of the presence of a linear element with the properties of an ideal filter in the system being investigated, posed by M. A. Aizerman [3], was considered earlier [4]. However, such an approximate solution does not give the capability of directly obtaining an answer to the question of taking into account the actual, nonidealized, frequency characteristic of the system's linear portion.

In the present paper we make the attempt to solve the problem exactly without having recourse to the assumption of the presence of an ideal filter.

The solution of the problem leads to the transformation of the characteristic equation of a system of equations with periodically varying coefficients, written in the form of an infinite determinant to a finite form, by using a method analogous to that employed in deriving the characteristic equation for the Hill equation.

Initially, we limit ourselves to the consideration of a system with many degrees of freedom with one nonlinear parameter, and we dwell on the transformation to finite form of the characteristic equation for the equation with small deviations from periodic motions. After the characteristic equation has been transformed to finite form, we can apply the well-known frequency criteria for analyzing stability. The methodology described is then generalized to the case of several equations with periodic coefficients.

The results obtained can be directly used in the solution of the enumerated problems.

Derivation of the Characteristic Equation

We consider an N-loop circuit in which the inductance of the principal arm of the ν'th loop is a nonlinear function of the current in the ν'th loop. In this circuit, let there exist a periodic mode with fundamental frequency Ω. For the given system, the equation for small deviations from the periodic motion can be presented in the following form. *

$$p\,[L_{\nu\nu}(t)\,i_\nu(t)] + \sum_{s=1}^{N} M_{\nu s}(p)\,i_s(t) = 0, \tag{1a}$$

$$\sum_{s=1}^{N} M_{\xi s}(p)\,i_\xi(t) = 0 \quad (\xi = 1, 2, \ldots, \nu-1, \nu+1, \ldots, N), \tag{1b}$$

where $M_{\nu s}$ and $M_{\xi s}$ are functions of the operator $p = d/dt$ of the form

$$M_{s\nu} = L_{s\nu}\,p + R_{s\nu} + \frac{1}{c_{s\nu}\,p}.$$

For the nonlinear inductance of the ν'th loop in the periodic mode we can write** [5]

$$L_{\nu\nu}(t) = 2l_{\nu\nu}^{(0)} + 2L_{\nu\nu}^{(1)}\cos\Omega t + \ldots + 2L_{\nu\nu}^{(m)}\cos m\,\Omega t + \ldots = \sum_{m=-\infty}^{\infty} L_{\nu\nu}^{m}\,e^{jm\Omega t}, \tag{2}$$

where $L_{\nu\nu}^{(0)} = 2l_{\nu\nu}^{(0)}$.

We will seek a solution of (1) in the form

$$i_\xi(t) = e^{\lambda t}\sum_{k=-\infty}^{\infty} I_\xi^{(k)}\,e^{j(k\Omega t+\varphi_k)} = e^{\lambda t}\sum_{k=-\infty}^{\infty} \dot{I}_\xi^{(k)}\,e^{jk\Omega t} \quad (\xi = 1, 2, \ldots, \nu, \ldots, N). \tag{3}$$

If, in (1b), we transfer all terms containing i_ν to the right side, we obtain a system of $(N-1)$ equations for the currents $i_s \to$ (for $s \neq \nu$). In system (1b) of equations, by expressing the currents i_s ($s \neq \nu$) in terms of i_ν, we get

$$i_s = Z_{s\nu}(p)\,i_\nu, \tag{4}$$

where $Z_{s\nu} = \dfrac{\Delta_s}{\Delta}$, $\Delta_s = \Delta_{1s} + \ldots + \Delta_{(n-1)s}$, Δ is the determinant of the system of equations obtained from (1b) after transferring i_ν to the right side, $\Delta_{\xi s}$ is the minor of determinant Δ obtained by striking out the ξ'th row and the s'th column.

By substituting the values of i_s thus found in (1a), we get

$$pL_{\nu\nu}(t)\,i_\nu(t) + \sum_{\substack{s=1\\ s\neq\nu}}^{N}(M_{\nu\nu} + M_{\nu s}Z_{s\nu})\,i_\nu(t) = 0 \tag{5}$$

*By using electrical engineering terminology, we can consider that the inductance or capacitance is found in the corresponding principal arm of the circuit.

**Here, as is usual, we assume that the sum $\sum_{m=0}^{\infty}|L_{\nu\nu}^{m}|$ is a finite quantity.

or

$$p\,[L_{\nu\nu}(t)\, i_\nu(t)] + Z_{\nu\,\mathrm{in}}\, i_\nu(t) = 0,$$

where $Z_{\nu\ \mathrm{in}}$ is the input impedance measured between the variable inductance terminals connected in the principal arm of the ν'th loop.

If, in (5), we substitute the values of $L_{\nu\nu}(t)$ and $i_\nu(t)$ from (2) and (3), we get

$$\sum_{k=-\infty}^{\infty} \left\{ (\lambda + jk\Omega) \left[L_{\nu\nu}^{(0)} + \frac{Z_{\nu\,\mathrm{in}}(\lambda + jk\Omega)}{\lambda + jk\Omega} \right] \dot{I}_\nu^{(k)}\, e^{(\lambda + jk\Omega)t} \right\} + \sum_{k=-\infty}^{\infty} \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} [\lambda + j(k+m)\,\Omega]\, L_{\nu\nu}^{(m)}\, \dot{I}_\nu^{(k)}\, e^{[\lambda + j(k+m)\Omega]t} = 0.$$

After replacing $(k + m)$ by $\underline{k}$ in the second sum, we get

$$\sum_{k=-\infty}^{\infty} \left\{ (\lambda + jk\Omega) \left[L_{\nu\nu}^{(0)} + \frac{Z_{\nu\,\mathrm{in}}(\lambda + jk\Omega)}{\lambda + jk\Omega} \right] \dot{I}_\nu^{(k)}\, e^{(\lambda + jk\Omega)t} \right\} + \sum_{k=-\infty}^{\infty} \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} (\lambda + jk\Omega)\, L_{\nu\nu}^{(m)}\, \dot{I}_\nu^{(k-m)}\, e^{(\lambda + jk\Omega)t} = 0. \tag{6}$$

If the coefficients of like powers of $\underline{e}$ are equated to zero, we obtain the system of equations

$$\left(L_{\nu\nu}^{(0)} + \frac{Z_{\nu\,\mathrm{in}}(\lambda + jk\Omega)}{\lambda + jk\Omega} \right) \dot{I}_\nu^{(k)} + \sum_{m=-\infty}^{\infty} L_{\nu\nu}^{(m)}\, \dot{I}_\nu^{(k-m)} = 0. \tag{7}$$

$$(k = -\infty, \ldots, -1, 0, 1, \ldots, \infty).$$

We now introduce the notation*

$$\frac{Z_{\nu\,\mathrm{in}}(\lambda + jk\Omega)}{\lambda + jk\Omega} = \frac{F_{\nu 1\,\mathrm{in}}(\lambda + jk\Omega)}{F_{\nu 2\,\mathrm{in}}(\lambda + jk\Omega)}. \tag{8}$$

By substituting (8) in (7) we get

$$[L_{\nu\nu}^{(0)} F_{\nu 2\,\mathrm{in}}(\lambda + jk\Omega) + F_{\nu 1\,\mathrm{in}}(\lambda + jk\Omega)]\, \dot{I}_\nu^{(k)} + F_{\nu 2\,\mathrm{in}}(\lambda + jk\Omega) \sum_{m=-\infty}^{\infty} L_{\nu\nu}^{(m)}\, \dot{I}_\nu^{(k-m)} = 0$$

or

$$\left[1 + \frac{L_{\nu\nu}^{(0)} F_{\nu 2\,\mathrm{in}}(\lambda + jk\Omega)}{F_{\nu 1\,\mathrm{in}}(\lambda + jk\Omega)} \right] \dot{I}_\nu^{(k)} + \frac{F_{\nu 2\,\mathrm{in}}(\lambda + jk\Omega)}{F_{\nu 1\,\mathrm{in}}(\lambda + jk\Omega)} \sum_{m=-\infty}^{\infty} L_{\nu\nu}^{(m)}\, \dot{I}_\nu^{(k-m)} = 0. \tag{9}$$

---

* Here, and below, we let $\dfrac{Z_{\nu\,\mathrm{in}}(\lambda + jk\Omega)}{\lambda + jk\Omega} \to \infty$ for $\mathrm{Re}\,|\lambda| \to \infty$.

We let $(\lambda + jk\Omega) = \mu_k$ and take $\alpha_1, \alpha_2, \ldots, \alpha_q$ as the roots of the polynomial $F_{\nu 1\,\mathrm{in}}(\lambda + jk\Omega) = F_{\nu 1\,\mathrm{in}}(\mu_k)$. Then, (9) can be rewritten in the form

$$\left\{1 + \sum_{i=1}^{n} \frac{A_{10}^{(i)}}{(\mu_k - \alpha_i)} + \sum_{i=1}^{n-r} \frac{A_{20}^{(i)}}{(\mu_k - \alpha_i)^2} + \sum_{i=1}^{n-r-t} \frac{A_{30}^{(i)}}{(\mu_k - \alpha_i)^3} + \ldots\right\} \dot{I}_{\nu}^{(k)} + \\ + \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \left\{\sum_{i=1}^{n} \frac{A_{1m}^{(i)}}{(\mu_k - \alpha_i)} + \sum_{i=1}^{n-r} \frac{A_{2m}^{(i)}}{(\mu_k - \alpha_i)^2} + \sum_{i=1}^{n-r-t} \frac{A_{3m}^{(i)}}{(\mu_k - \alpha_i)^3} + \ldots\right\} \dot{I}_{\nu}^{(k-m)} = 0, \tag{10}$$

where $\underline{n}$ is the number of different roots of polynomial $F_{\nu 1\,\mathrm{in}}(\mu_k)$, $\underline{r}$ is the number of simple roots, $\underline{t}$ is the number of double roots, etc.

We now write the determinant of system (10), equal to $\Delta(\lambda)$.

By substituting $\mu_k = \lambda + jk\Omega$, we transform the determinant $\Delta(\lambda)$.

Without computing the determinant $\Delta(\lambda)$, we limit ourselves to writing, in general form, its diagonal and nondiagonal elements. For this we shall initially assume, for simplicity though without destroying the generality of the reasoning, that all the poles are either simple or double. A diagonal term of this determinant will have the following form:

$$1 + \sum_{i=1}^{n} \frac{A_{10}^{(i)}}{(\lambda + jk\Omega - \alpha_i)} + \sum_{i=1}^{n-r} \frac{A_{20}^{(i)}}{(\lambda + jk\Omega - \alpha_i)^2}; \tag{11}$$

and a nondiagonal term will equal

$$\sum_{i=1}^{n} \frac{A_{1m}^{(i)}}{\lambda + jk\Omega - \alpha_i} + \sum_{i=1}^{n-r} \frac{A_{2m}^{(i)}}{(\lambda + jk\Omega - \alpha_i)^2}. \tag{12}$$

Thus, the difference between the determinant obtained here and the corresponding determinant for the Hill equation is defined by the number of poles and their multiplicity which, in the general case, here, can have arbitrary values, and by the presence of the term α in the denominators of the determinant's elements.

The determinant $\Delta(\lambda)$ is a meromorphic function with poles at the points $(\alpha_1 - jk\Omega;\ \alpha_2 - jk\Omega;\ \alpha_n - jk\Omega;\ \ldots)$. If we expand this determinant in a Laurent series in the neighborhood of the point $\alpha_1 - jk\Omega$ we obtain

$$\Delta(\lambda) = \Delta_1(\lambda) + \Delta_2(\lambda) = \Delta_1(\lambda) + \sum_{i=0}^{n} \Delta_{2i}^{(1)} + \sum_{i=0}^{n-r} \Delta_{2i}^{(2)}, \tag{13}$$

where $\Delta_1(\lambda)$ is the integral part and $\Delta_2(\lambda)$ is the fractional part of the Laurent series expansion which, in the case of a simple pole, equals

$$\Delta_{2i}^{(1)}(\lambda) = \sum_{k=-\infty}^{\infty} C_{ik}^{(1)} (\lambda - \alpha_1 + jk\Omega)^{-1} \quad (i = 1, 2, \ldots, n) \tag{14}$$

and, in the case of a double pole, equals

$$\Delta_{2i}^{(2)}(\lambda) = \sum_{k=-\infty}^{\infty} C_{ik}^{(2)} (\lambda - \alpha_i + jk\Omega)^{-2} \quad (i = 1, 2, \ldots, n - r). \tag{15}$$

The coefficient $C_{ik}^{(1)}$ equals (in the case of a simple pole*

$$C_{ik}^{(1)} = \lim_{\lambda \to (\alpha_i - jk\Omega)} [\Delta(\lambda)(\lambda - \alpha_i + jk\Omega)] \quad (16)$$

and for a multiple pole of multiplicity two

$$C_{ik}^{(2)} = \lim_{\lambda \to (\alpha i - jk\Omega)} [\Delta(\lambda)(\lambda - \alpha_i + jk\Omega)^2]. \quad (17)$$

We take into account that, for all values of k, $C_{ik} = D_i$, where D_i is the determinant analogous to the Hill determinant for the case of a simple pole, equal to $D_i^{(1)}$.

The elements of the center row of this determinant and of the row shifted to k rows below the center row will have, respectively, the forms

$$\ldots; \quad \frac{A_{i,1}}{1}; \quad \frac{A_{i,0}}{1}; \quad \frac{A_{i,-1}}{1}; \quad \ldots$$

and

$$\ldots; \sum_{\xi=1}^{n} \frac{A_{i,1}}{\alpha_i - \alpha_\xi + jk\Omega}; \; 1 + \sum_{\xi=1}^{n} \frac{A_{i,0}}{\alpha_i - \alpha_\xi + jk\Omega} + \ldots; \; \sum_{\xi=1}^{n} \frac{A_{i,-1}}{\alpha_i - \alpha_\xi + jk\Omega}; \ldots$$

(for double poles, the denominators of the determinant elements must be raised to the second power, etc.).

The sum of all the component fractional parts of the Laurent series of the function $\Delta(\lambda)$, corresponding to a given subscript i, for the case of a double pole equals [6]

$$\sum_{i=1}^{n-r} \Delta_{2i}^{(2)}(\lambda) = \sum_{i=1}^{n-r} - \frac{D_i^{(2)} \pi^2}{\Omega^2 \sin^2 \left[(\lambda - \alpha_i) \pi \frac{1}{j\Omega} \right]}. \quad (19)$$

Correspondingly, the component sum of the fractional parts of the function for the subscript i in the case of a simple pole will equal**

$$\sum_{i=1} \Delta_{2i}^{(1)}(\lambda) = \sum_{i=1}^{n} \Delta_{2i}^{(1)}(\lambda) = \sum_{i=1}^{n} \operatorname{ctg} \left[\pi (\lambda - \alpha_i) \frac{1}{j\Omega} \right] D_i^{(1)} \frac{\pi}{j\Omega}. \quad (20)$$

By using expressions (19-20) we can obtain the characteristic equation in finite form.

---

*The coefficient C_{ik}^{r}, corresponding to a simple fraction expansion, is found by well-known rules for multiple poles. For example, the coefficient $C_{ik}^{(t)}$ from a second-order pole will equal

$$C_{ik}^{(t)} = \left\{ \frac{d}{d\lambda} [\Delta(\lambda)(\lambda - \alpha_i + jk\Omega)^2] \right\}_{\lambda = \alpha_i - jk\Omega}.$$

**By taking into account that d cot $\nu / d\nu = -\sin^2 \nu$, if there are poles of third and higher order in Z(p) we will find, in the expression for $\Delta(\lambda)$, linear combinations of the functions $\frac{\pi^3}{j\Omega} \frac{d^2}{d\lambda^2} \left[\frac{\pi}{j\Omega} (\lambda - \alpha_i) \right]$ etc.

In correspondence with (11-14) we can write

$$\Delta(\lambda)=\Delta_1(\lambda)+\Delta_2(\lambda)=\sum_{i=1}^{n}\Delta_{1i}(\lambda)+\sum_{i=1}^{n}\Delta_{2i}^{(1)}(\lambda)+\sum_{i=1}^{n-r}\Delta_{2i}^{(1)}(\lambda). \tag{21}$$

In correspondence with (19-20), as $\mathrm{Re}\lambda\to\infty$, $\Delta_2(\lambda)$ is a bounded quantity.

We note that the function $\Delta(\lambda)$ has period $j\Omega$, in correspondence with which we limit ourselves to a consideration of the function in the band $|\mathrm{Im}\,\lambda|\leqslant\frac{\Omega}{2}$. With this, $\lim\limits_{\mathrm{Re}\,\lambda\to\infty}\Delta_2(\lambda)=0$, a $\lim\limits_{\mathrm{Re}\lambda\to\infty}\Delta_1(\lambda)=\lim\limits_{\mathrm{Re}\lambda\to\infty}\Delta(\lambda)$ has a finite value equal to unity, i.e.,

$$\sum_{i=1}^{n}\Delta_{1i}(\lambda)=1.$$

In the particular case of the Hill equation, we obtain from (19) and (21) that

$$\sin^2\left[\pi(\lambda-\alpha_i)\frac{1}{j\Omega}\right]=\pi^2 D_1^{(2)}\frac{1}{\Omega^2}. \tag{22}$$

After the characteristic equation has been put in finite form, one can use one of the well-known frequency criteria for stability analysis. It should be taken into account that the use of frequency criteria has its own peculiar features here. These peculiar features are determined by the fact that, in the given case, the characteristic function is a periodic function of frequency. In correspondence with this, it suffices here, as in the analogous case of pulse-amplitude modulation, to construct the hodograph of $\Delta(j\omega)$ as ω varies from $-\Omega/2$ to $\Omega/2$.

We consider briefly the application of the Mikhailov frequency criterion to the system of equations of motion for which the characteristic equation has the form of (22). If α_1 is a pure imaginary number, then the Mikhailov hodograph of the left member of equation (22) will be a segment of a line coinciding with the real axis.

Correspondingly, if α_1 is a real or complex number, the Mikhailov hodograph of the left member of (22) will be an infinite number of ellipses, superimposed one on another, and passing by turns through all quadrants. For small values of the left member of (22), i.e., for small values of $D_1^{(2)}$, the Mikhailov hodograph will pass by turns through all quadrants, and the system will be stable. As the determinant $D_1^{(2)}$ increases, the right member of (22) increases, and the ellipses corresponding to the left member of (22) are shifted to the right and, for a certain limiting value of determinant $D_1^{(2)}$, the system becomes unstable. The limiting value of determinant $D_1^{(2)}$, corresponding to the boundary of stability, is found from the condition [3]

$$\mathrm{Re}\left[\sin^2(j\omega_1-\alpha_1)\frac{\pi}{j\Omega}\right]=D_1^{(2)}\pi^2\frac{1}{\Omega^2}, \tag{23}$$

where $D_1^{(2)}$ depends on the modulation depth and ω_1 is the most negative root of the equation

$$I_m\sin^2\left[(j\omega-\alpha_1)\frac{\pi}{j\Omega}\right].$$

On the basis of what has been presented here, we can solve the problem of contructing a control system which satisfies the conditions of stability for some periodic mode with fundamental frequency Ω. As an example, we consider the control system described by the equations

$$D(p)X=-K(p)Y,\quad Y=F(X). \tag{24}$$

By means of a transformation analogous to those given above [3, 4], we arrive at the following infinite system of equations:

$$
\begin{gathered}
\{[A_0K(p)+D(p)]\,e^{(\lambda+jk\Omega)t}\}\,\dot{X}_k+\left[\frac{1}{2}A_0m_1K(p)\,e^{(\lambda+jk\Omega)t}\right]\dot{X}_{k-1}+\\
+\left[\frac{1}{2}A_0m_1K(p)\,e^{(\lambda+jk\Omega)t}\right]\dot{X}_{k+1}+\ldots+\left[\frac{1}{2}A_0m_rK(p)\,e^{(\lambda+jk\Omega)}\right]\dot{X}_{k-r}+\\
+\left[\frac{1}{2}A_0m_rK(p)\,e^{(\lambda+jk\Omega)t}\right]\dot{X}_{k+r}+\ldots=0,
\end{gathered}
\tag{25}
$$

where A_0m_k is the coefficient of the k'th harmonic in the Fourier series expansion of the function

$$
A_k(t)=\left[\frac{\partial F(X)}{\partial X}\right]_{X=X(t)}.
$$

By setting up the determinant of system (25) and equating it to zero, we obtain the characteristic equation of the system.

By dividing the k'th row of the determinant thus obtained by D(p), we obtain the determinant with diagonal terms of the form

$$
1+A_0\frac{K(\lambda+j\Omega)}{D(\lambda+j\Omega)}
$$

and nondiagonal terms of the form

$$
\frac{1}{2}A_0m_r\frac{K(\lambda+j\Omega)}{D(\lambda+j\Omega)}.
$$

We make the analogous assumption that

$$
\frac{K(\lambda+j\Omega)}{D(\lambda+j\Omega)}\to 0\ \text{ for }\ Re(\lambda)\to\infty.
$$

Let the multinomial D(p) have n different roots, of which r are simple and t double, and then the determinant of system (25) can be given in a form analogous to that of the determinant of system (10).

By transforming the infinite determinant thus obtained to a finite form in correspondence with (14-20), just as this was done in the case considered above, we obtain the characteristic equation in the form

$$
\sum_{i=1}^{n}\cot\left[\pi(\lambda-\alpha_i)\frac{1}{j\Omega}\right]D_i^{(1)}\frac{1}{j\Omega}-\sum_{i=1}^{n-r}\frac{D_i^{(2)}\pi^2}{\Omega^2\sin^2\left[(\lambda-\alpha_i)\pi\frac{1}{j\Omega}\right]}=1.
$$

Here, the difference from the previous case is determined only by the denominators of the elements of determinant D.

In conclusion, we consider the more general case when the variable parameters of L(t) occur in b loops, where b < n. In this case, instead of system (1) of equations, we can write

$$
\begin{gathered}
p\,[L_{\nu\nu}(t)\,i_\nu(t)]+\sum_{s=1}^{b}M_{\nu s}(p)\,i_s(t)+\sum_{\xi=b+1}^{N}M_{\nu\xi}(p)\,i_\xi(t)=0\quad(\nu=1,2,\ldots,b),\\
\sum_{s=1}^{n}M_{\xi s}(p)\,i_\xi(t)=0\quad(\xi=b,\ldots,n)
\end{gathered}
\tag{26}
$$

where N is the number of loops, ν is the number of a loop containing the variable inductance, ξ is the number of a loop containing only invariable parameters.

In contradistinction to the cases considered above, we here assume that the periodic function $L_{\nu\nu}(t)$ can be given in the form of a finite series, containing m_1 terms. The solution, as above, will be sought in the form of (3).

If we express all the i_ξ in terms of the i_ν, and carry out all the transformations given above in the transition from (1) to (6), we obtain a system of <u>b</u> equations in the form

$$\begin{gathered}\sum_{k=-\infty}^{\infty}(\lambda+jk\Omega)\left[L_{\nu\nu}^{(0)}+\frac{Z_\nu\,\text{in}\,(\lambda+jk\Omega)}{\lambda+jk\Omega}\right]\dot{I}_\nu^{(k)}e^{(\lambda+jk\Omega)t}+\\+\sum_{k=-\infty}^{\infty}\sum_{\substack{m=-m_1\\m\neq 0}}^{m_1}[\lambda+j(k+m)\Omega]\,L_{\nu\nu}^{(m)}\dot{I}_\nu^{(k)}e^{[\lambda+j(k+m)\Omega]t}+\\+\sum_{k=-\infty}^{\infty}\sum_{\substack{s=1\\s\neq\nu}}^{b}(\lambda+jk\Omega)\,M_{\nu s}(\lambda+jk\Omega)\,\dot{I}_s^{(k)}e^{(\lambda+jk\Omega)t}=0\quad(\nu=1,2,\ldots,b).\end{gathered}\tag{27}$$

On the basis of (27) we can, analogously to what was done previously, write

$$\begin{gathered}(\lambda+jk\Omega)\left[L_{\nu\nu}^{(0)}+\frac{Z_\nu\,\text{in}(\lambda+jk\Omega)}{\lambda+jk\Omega}\right]\dot{I}_\nu^{(k)}+\sum_{m=-m_1}^{m_1}(\lambda+jk\Omega)\,L_{\nu\nu}^{(m)}\,\dot{I}_\nu^{(k-m)}+\\+\sum_{\substack{s=1\\s\neq\nu}}^{b}(\lambda+jk\Omega)\,M_{\nu s}(\lambda+jk\Omega)\,\dot{I}_s^{(k)}=0\\(k=-\infty,\ldots,-1,0,1,\ldots,\infty),\quad(\nu=1,2,\ldots,b).\end{gathered}\tag{28}$$

System (28) of equations, for each value, contains <u>b</u> equations in $b + 2m_1$ unknowns. In this system, the unknowns are the variables

$$\dot{I}_\nu^{(k)},\ \dot{I}_\nu^{(k+1)},\ \dot{I}_\nu^{(k-1)},\ldots,\dot{I}_\nu^{(k+m_1)},\ \dot{I}_\nu^{(k-m_1)},\ldots(\nu=1,2,\ldots,b).$$

By using a method, analogous to [4], of replacing <u>k</u> by (k + 1), (k −1), etc., we increase the system of equations to $b + 2m_1$ equations, after which we eliminate from them all variables, $\dot{I}_\nu^{(k)}$, $\dot{I}_\nu^{(k+m)}$ etc., for all ν except, for example, $\nu = 1$. As a result, we obtain expressions for $\dot{I}_1^{(k)}$, $\dot{I}_1^{(k+1)}$, . . . which are analogous to (7), after which the necessary result can be obtained by repeated carrying out of the above operation.

The author wishes to express his gratitude to I. M. Smirnova for reading the manuscript and making a number of valuable remarks.

Received March 6, 1959

LITERATURE CITED

[1] L. S. Gol'dfarb, "On certain nonlinearities in automatic control systems," [in Russian]. Automation and Remote Control (USSR) 8, 5 (1947).

[2] M. A. Aizerman and I. M. Smirnova, "On applying the method of small parameters in investigating periodic modes in automatic control systems," [in Russian]. Collected Papers in Memory of A. A. Andronov. Izdatelstvo AN SSSR (Moscow, 1955).

[3] M. A. Aizerman, Lectures on Automatic Control Theory [in Russian]. GTTI (Moscow, 1956, 1958).

[4] V. A. Taft, "On the stability of periodic modes in nonlinear automatic control systems containing linear filters," [in Russian]. Automation and Remote Control (USSR) 19, 7 (1958).*

[5] V. A. Taft, Electric Circuits with Periodically Varying Parameters and Transient Responses in Synchronous Machines, [in Russian]. Izdatelstvo AN SSSR (1958).

[6] Z. B. Ryzhik and I. S. Gradshtein, Tables of Sums, Integrals and Products [in Russian] GTTI (Moscow, 1954).

* See English translation.

COMPUTING THE TRANSIENT RESPONSE IN LINEAR SYSTEMS BY THE METHOD OF REDUCING THE ORDER OF THE DIFFERENTIAL EQUATION

A. V. Kalyaev
(Taganrog)

The paper considers the computation of the transient responses in linear systems with constant parameters by the method of reducing the order of the differential equation. Examples are given of the application of the method developed.

The question of calculating the nonstationary-state modes in linear systems has been developed in great detail. There exist such powerful instruments for analyzing these processes as the operator method. Its disadvantage is the necessity of determining the roots of the characteristic equation which, for high orders of the differential equation, is a quite difficult task which can only be solved in numerical form. In order to avoid this difficulty a significant number of methods were suggested for the approximate construction of the transient responses in linear systems which allow the computation of the transient responses to be carried out without seeking the roots of the characteristic equation [1-4]. The development of computers, both analog and digital, allowed one to analyze, quite rapidly and simply, the transient responses both in linear and in nonlinear dynamic systems.

The mathematic models and digital machines of today forced the approximate methods for computing transient responses into the background. The value of approximate methods today is retained only in so far as they can be used for setting up the simplest algorithm for solving the problem which is necessary for programming.

However, it might still be urged that the value of approximate methods of analysis is far from being exhausted. Indeed, neither mathematical machines nor the existing contemporary methods of approximate calculation permit the solution of the problem being considered in a general literal form. It is unnecessary to prove that the importance of obtaining such a solution is in no way diminished as a result of the vigorous development of computational technology. If one succeeds in analyzing, at least approximately, the transient responses in high-order dynamic systems in general literal form, then this will be an important step in the development of the theory of nonsteady-state modes.

In this paper we consider the question of seeking an approximate solution of a linear differential equation with constant coefficients by reducing the order of the equation, which in many cases permits one to obtain the solution in the most general literal form.

We assume that there is some dynamic system which, within certain limits of its behavior, can be described by a linear differential equation with constant coefficients

$$\sum_{k=0}^{k=n} c_k y^{(k)} = f(t). \tag{1}$$

In the majority of cases one can set up a sufficiently accurate representation of the properties of the linear system to be analyzed if one allows to act on the latter, not a signal of arbitrary form, but one of the standard signals. In particular, for analysing linear systems, wide usage is made of a standard signal in the form of a unit function, $f(t) = 1(t)$. In this case, equation (1) takes the following form:

$$\sum_{k=0}^{k=n} c_k y^{(k)} = 1(t). \tag{2}$$

If the right member of the linear differential equation is a constant, then a simple transformation of the (y, t) coordinate system, translating it along the y axis, will take equation (2) into an equation with a zero right member. In the sequel we shall assume that such a coordinate transformation has been carried out and that we are given the following equation to be analyzed

$$\sum_{k=0}^{k=n} c_k y^{(k)} = 0 \tag{3}$$

with given initial conditions (for $t = 0$):

$$y^{(k)} = y_0^{(k)} \quad [k = 0, 1, 2, \ldots, (n-1)]. \tag{4}$$

We limit ourselves to the investigation of stable systems, the transient responses in which play an important role in practical cases.

If there are no multiple roots, the solution of equation (3) has, as is well-known, the form:

$$y(t) = \sum_{k=1}^{k=n} a_k e^{\alpha_k t}. \tag{5}$$

The presence of multiple roots, for which the solution of equation (3) is written in the form

$$y(t) = \sum_k P_{\gamma_k - 1}(t) e^{\alpha_k t},$$

where $P_{\gamma_k - 1}(t)$ is a polynomial of degree $(\gamma_k - 1)$ in the variable t, gives nothing new to the theory and will not enter into the discussion which follows. Thus, we shall consider below only the solution in cases where there are no repeated roots.

It frequently turns out that some of the terms of expression (5) are either small in absolute magnitude or are rapidly damped. Under these conditions it becomes possible to lower the order of differential equation (3) and, thereby, to simplify its solution.

In fact, it is possible to approximate the exact curve y(t) by some approximate function

$$y_{ap}(t) = \sum_{k=1}^{k=m} a_k e^{\alpha_k t}, \tag{6}$$

for which m < n, i.e., the number of terms in the right member of expression (6) is less than the number of terms in the right member of expression (5). On the other hand, expression (6) corresponds to some differential equation

$$\sum_{k=0}^{k=m} b_k y_{\mathrm{ap}}^{(k)} = 0, \tag{7}$$

whose order is lower than the order of the original equation (3).

We assume for definiteness that the initial conditions which correspond to equation (7) are the same as in the case of equation (3) (for t = 0):

$$y_{\mathrm{ap}}^{(k)} = y_0^{(k)} \quad [k = 0, 1, 2, \ldots, (m-1)]. \tag{8}$$

The coefficients b_k (k = 0, 1, 2, . . ., m) are unknown. If we succeed in determining them by some means or other, then we can seek, not the exact solution of equation (3), but the approximate solution, $y_{\mathrm{ap}}(t) \approx y(t)$, the exact solution of equation (7) which has a lower order than the initial equation (3). Thus, we set ourselves the problem of determining the coefficients b_k (k = 0, 1, 2, . . ., m). Without loss of generality we may choose one of these coefficients arbitrarily. For simplicity, we set $b_m = 1$. The remaining coefficients must be so chosen that the approximate function $y_{\mathrm{ap}}(t)$ differs as little as possible from the exact function y(t). We require, for this, that the error which arises as the result of substituting the exact function y(t) in the approximating equation (7) be minimal:

$$\delta(t) = \sum_{k=0}^{k=m} b_k y^{(k)}(t). \tag{9}$$

We now set up the integral of the squared error

$$I_\delta = \int_0^\infty \delta^2(t)\,dt = \int_0^\infty \Big[\sum_{k=0}^{k=m} b_k y^{(k)}(t)\Big]^2 dt, \tag{10}$$

which has a minimum when the following conditions hold

$$\frac{\partial I_\delta}{\partial b_i} = 0 \quad [i = 0, 1, 2 \ldots, (m-1)]. \tag{11}$$

By substituting the value of I_δ from (10) in (11) and differentiating, we obtain the following system of equations:

$$\sum_{k=0}^{k=m-1} b_k \int_0^\infty y^{(k)} y^{(i)}\,dt = -\int_0^\infty y^{(m)} y^{(i)}\,dt \quad [i = 0, 1, 2, \ldots, (m-1)], \tag{12}$$

which allows us to determine the coefficients b_k. The coefficients obtained lead to a minimum of the error, $\delta(t)$, but do not provide a minimum of the error

$$\varepsilon(t) = y(t) - y_{\mathrm{ap}}(t). \tag{13}$$

To decrease the error $\epsilon(t)$ one can require, in addition, that the following condition hold

$$I_\epsilon = \int_0^\infty \varepsilon(t)\,dt = \int_0^\infty [y(t) - y_{ap}(t)]\,dt = 0, \tag{14}$$

while simultaneously discarding one of equations (12).

It follows from relationship (14) that

$$\int_0^\infty y_{ap}(t)\,dt = \int_0^\infty y(t)\,dt. \tag{15}$$

The integral $\int_0^\infty y_{ap}(t)\,dt$ can be determined by using equation (7). Indeed, by integrating this equation between the limits of 0 and ∞, we obtain

$$\sum_{k=0}^{k=m} b_k \int_0^\infty y_{ap}^{(k)}(t)\,dt = 0,$$

from whence, taking the initial conditions in (8) into account, we find that

$$b_0 \int_0^\infty y_{ap}(t)\,dt = \sum_{k=1}^{k=m} b_k y_0^{(k-1)}.$$

Thus, by taking (15) into account, we can write

$$b_0 \int_0^\infty y(t)\,dt = \sum_{k=1}^{k=m} b_k y_0^{(k-1)}. \tag{16}$$

As the result, for determining the coefficients b_k [$k = 0, 1, 2, \ldots, (m-1)$], we obtain the following system of linear algebraic equations:

$$\begin{gathered} \sum_{k=0}^{k=m-1} b_k \int_0^\infty y^{(k)} y^{(i)}\,dt = -\int_0^\infty y^{(m)} y^{(i)}\,dt \quad [i = 0, 1, 2, \ldots (m-2)], \\ \sum_{k=0}^{k=m-1} b_k y_0^{(k-1)} = -y_0^{(m-1)}, \end{gathered} \tag{17}$$

where we have used the notation

$$y_0^{(-1)} = -\int_0^\infty y(t)\,dt.$$

The constant quantities of the form $\int_0^\infty y(t)\,dt$ and $\int_0^\infty y^{(k)} y^{(i)}\,dt$, which enter into this system may be easily determined by using the initial differential equation (3) and the corresponding initial conditions (4). Indeed, if we multiply equation (3) termwise by the quantities $y^{(i)}$ [$i = 0, 1, 2, \ldots, (n-1)$] and integrate between the limits of 0 and ∞ while taking initial conditions (4) and final conditions $y_\infty^{(i)} = 0$ [$i = 0, 1, 2, \ldots (n-1)$] into account, we obtain the system of algebraic equations which is necessary for computing integrals of the type $\int_0^\infty y^{(k)} y^{(i)}\,dt$:

$$\sum_{k=0}^{k=n} c_k \int_0^\infty y^{(k)} y^{(i)}\,dt = 0 \qquad [i = 0, 1, 2, \ldots, (n-1)]. \tag{18}$$

This system of $\underline{n}$ linear algebraic equations contains $[n(n+3)]/2$ definite integrals:

$$Y_{ki} = \int_0^\infty y^{(k)} y^{(i)}\,dt \qquad \begin{pmatrix} k = 0, 1, 2, \ldots, n, \\ i = 0, 1, 2, \ldots, n-1 \end{pmatrix}.$$

However, in actuality, only $\underline{n}$ integrals of the following type are unknown

$$Y_{kk} = \int_0^\infty [y^{(k)}]^2\,dt \qquad [k = 0, 1, 2, \ldots, (n-1)]. \tag{19}$$

The remaining $[n(n+1)]/2$ integrals are either removed by integrating by parts, if the sum $(k+i)$ is odd:

$$\int_0^\infty y^{(k)} y^{(i)}\,dt = \Big[\sum_{p=0}^{p=\frac{1}{2}(k-i-3)} (-1)^p y^{(k-p-1)} y^{(i+p)}\Big]_{t=0}^{t=\infty} + \\ + \Big\{\frac{(-1)^{\frac{1}{2}(k-i-1)}}{2}[y^{\frac{1}{2}(k+i-1)}]^2\Big\}_{t=0}^{t=\infty} \quad (k > i),$$

or are reduced to integrals of the form of (19) if the sum $(k+i)$ is even:

$$\int_0^\infty y^{(k)} y^{(i)}\,dt = \Big[\sum_{p=0}^{p=\frac{1}{2}(k-i-2)} (-1)^p y^{(k-p-1)} y^{(i+p)}\Big]_{t=0}^{t=\infty} + \\ + (-1)^{\frac{1}{2}(k-i)} \int_0^\infty [y^{\frac{1}{2}(k+i)}]^2\,dt \quad (k > i).$$

We note that system (18) of equations obtained above is compatible, since its determinant, as is easily shown, reduces to the Hurwitz determinant for Equation (3) and, consequently, is different from zero.

The integral $\int_0^\infty y(t)\,dt$, which enters into system (17) of equations is simply computed by integrating equation (4) between the limits of zero and infinity:

$$\int_0^\infty y(t)\,dt = \sum_{k=1}^{k=n} \frac{c_k}{c_0} y_0^{(k-1)}. \qquad (20)$$

Thus, by computing integrals of the types $\int_0^\infty y\,dt$ and $\int_0^\infty y^{(k)} y^{(i)}\,dt$ by means of expressions (18) and (20) and then, from the system (17) of linear algebraic equations, determining the coefficients b_k [$k = 0, 1, 2, \ldots, (m-1)$], we take the problem of analyzing the transient response in the system described by equation (3) and transform it to the simpler problem of determining its approximate solution by means of a differential equation of lower order, namely, equation (7). The size of the error engendered by this can be estimated by computing the integral

$$I_\delta = \int_0^\infty \delta^2(t)\,dt,$$

which presents no difficulties in theory if the approximate solution, $y_{ap}(t)$, is determined. Obviously, for the very best approximating function $y_{ap}(t)$ to the exact function $y(t)$, it is necessary that the following inequality hold

$$I_\delta = \int_0^\infty \delta^2(t)\,dt \ll \int_0^\infty y^2(d)\,dt,$$

where the integral $\int_0^\infty y^2(t)\,dt$ is determined from system (18) of equations.

The results of carrying out the calculations given here for concrete linear dynamic systems operating in the nonsteady-state modes support the correctness of the method presented of lowering the order of linear differential equations with constant coefficients.

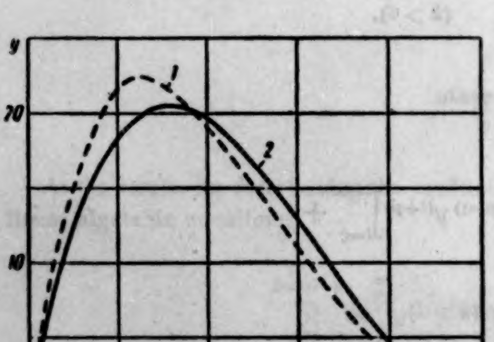


Fig. 1. 1) is the exact computed curve; 2) is the approximate solution; $y = 47.6\,e^{-0.281\,t}\sin 0.42\,t$.

In particular, the transient responses in several actual automatic control systems were computed by means of the method herein developed (Appendix I and II). With this, the order of the original differential equation was reduced by a factor or two or three. As the results of these computations, shown in Figs. 1 and 2, demonstrate, the approximate curve, $y_{ap}(t)$, does not differ much from the corresponding exact curve.

It should be emphasized that all the transformations carried out in decreasing the order of an equation are linear ones and, therefore, may be implemented in general literal form. With this, if the lowering is carried out down to the second order then, in certain cases, it is completely possible that the problem of investigating the transient response will be solved in general literal form. As an example of such an analysis of transient responses, Appendix III gives the computation of the transient response in a third-order system. Naturally, the method presented is not universal and, in certain cases, may lead to significant errors, but

for the most widely-used automatic control systems it is completely effective and can give satisfactory results. In all cases the method given can be recommended for investigating the quality of the transient response. The

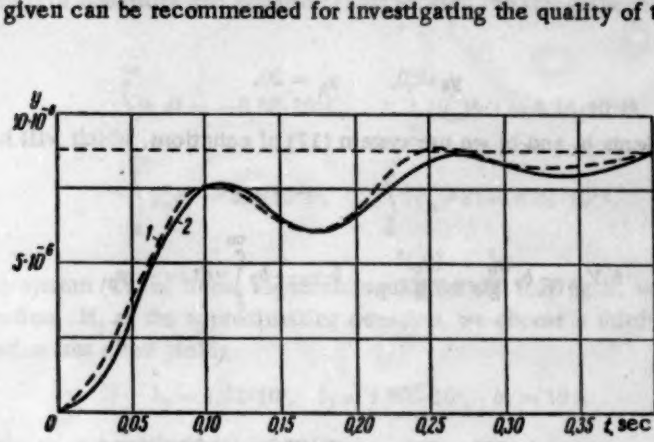


Fig. 2. 1) is the exact computed curve; 2) is the approximate solution, $y = [8.78 - 6.67\,e^{-8.06t} - 3.0\,e^{-5.72t}\sin(38.5t + 0.755)]\,10^{-6}$.

basic quality indicators – the duration of the transient response, the size of the overshoot and the number of oscillations – may, by this method, be computed with sufficient accuracy.

Appendix I

As an example, we consider the computation of the transient response in the system "airplane with autopilot" [1]. Such a system can be described by a sixth-order linear differential equation:

$$y^{(6)} + 16.4y^{(5)} + 107.4y^{(4)} + 364.2y^{(3)} + 1146.5y'' + 771.2y' + 292.1y = 0 \tag{21}$$

We assume that the initial conditions are given (for t = 0):

$$y_0 = 0, \quad y_0' = 20, \quad y_0'' = 0, \quad y_0^{(3)} = 0, \quad y_0^{(4)} = 0, \quad y_0^{(5)} = 0. \tag{22}$$

The exact solution of equation (21), obtained in [1], looks like:

$$y(t) = 0.164e^{-7.15t}\cos(3.35t + 0.732) - 1.232e^{-0.648t}\cos(3.74t - 1.312) + \\ + 59.6e^{-0.377t}\cos(0.428t + 1.57). \tag{23}$$

We now determine the approximate solution of differential equation (21) by the method of lowering its order. By using system (18) of equations, set up for original equation (21) and initial conditions (22), we determine the integrals:

$$Y_{00} = \int_0^\infty y^2 dt = 1076, \quad Y_{11} = \int_0^\infty (y')^2 dt = 356, \quad Y_{22} = \int_0^\infty (y'')^2 dt = 250. \tag{24}$$

Equation (20) allows us to compute the integral:

$$\int_0^\infty y(t)\,dt = 78.5. \tag{25}$$

We approximate original equation (21) by the second-order linear differential equation

$$y'' + b_1 y' + b_0 y = 0, \tag{26}$$

where, for t = 0,

$$y_0 = 0, \quad y'_0 = 20.$$

To compute the coefficients b_0 and b_1 we use system (17) of equations, which will have the following form in the given case:

$$b_1 Y_{11} - b_0 \frac{y_0^2}{2} = \frac{(y'_0)^2}{2}, \quad b_1 y_0 - b_0 \int_0^\infty y\,dt = -y'_0. \tag{27}$$

From this we have

$$b_1 .356 = 200, \quad -b_0 \cdot 78.5 = -20$$

or

$$b_1 = 0.562, \quad b_0 = 0.255.$$

Consequently, the approximating equation looks like:

$$y'' + 0.562y' + 0.255y = 0. \tag{28}$$

Finally, we determine its solution with the given initial conditions

$$y = 47.6e^{-0.281t} \sin 0.42t. \tag{29}$$

The curves corresponding to the exact solution, (23), of initial equation (21) and to its approximate solution, (29), are given in Fig. 1. As follows from the figure, despite the lowering of the differential equation's order by a factor of three, the solution of the approximating differential equation (28) coincides quite satisfactorily with the accurate solution, (23), of initial differential equation (21).

Appendix II

We now consider still another example of the computation of the transient response by the method of lowering the order of the differential equation, this time for an electro-hydraulic servo system [1] described by the equation

$$y^{(4)} + 103y^{(3)} + 3065y'' + 149250y' + 1081500y = 9.5, \tag{30}$$

where, for t = 0,

$$y_0 = 0, \quad y'_0 = 0, \quad y''_0 = 3.5 \cdot 10^{-8}, \quad y'''_0 = 9.5 \cdot 10^{-3}.$$

In the steady-state, we obviously have

$$y_s = \frac{9.5}{1081500} = 8.78 \cdot 10^{-6}.$$

Consequently, for the transient component, $y_n(t)$, of the function sought, $y(t)$, taking the relationship $y(t) = y_n(t) + y_s(t)$ into account, we obtain the differential equation:

$$y_n^{(4)} + 103y_n^{(3)} + 3065y''_n + 149250y'_n + 1081500y_n = 0 \tag{31}$$

with initial conditions (for t = 0)

$$y_{n0} = -8.78 \cdot 10^{-6} \quad y'_{n0} = 0, \quad y''_{n0} = 3.5 \cdot 10^{-8}, \quad y'''_{n0} = 9.5 \cdot 10^{-3}.$$

By using system (18) of equations and equation (20), we find the integrals:

$$\int_0^\infty y_n dt = -8.68\cdot 10^{-7}, \qquad \int_0^\infty (y_n')^2 dt = 8.16\cdot 10^{-10},$$
$$\int_0^\infty y_n^2 dt = 3.9\cdot 10^{-13}, \qquad \int_0^\infty (y_n'')^2 dt = 8.34\cdot 10^{-7}. \tag{32}$$

By now setting up system (17) of linear algebraic equations and solving it, we determine the coefficients of the approximating equation. If, as the approximating equation, we choose a third-order linear differential equation, then the computation just cited yields

$$b_0 = 1.22\cdot 10^4, \quad b_1 = 1.605\cdot 10^3, \quad b_2 = 19.5.$$

Thus, the approximating equation will look like:

$$y_n''' + 19.5y_n'' + 1605y_n' + 12200y_n = 0; \tag{33}$$

for t = 0

$$y_{n0} = 8.78\cdot 10^{-6}, \quad y_{n0}' = 0, \quad y_{n0}'' = 3.5\cdot 10^{-3}.$$

Its solution has the following form:

$$y_n = -6.67\cdot 10^{-6}e^{-8.06t} - 3.09\cdot 10^{-6}e^{-5.72t}\sin(38.5t + 0.755). \tag{34}$$

Hence, the approximate expression for the function sought, $y = y_n + y_s$ is:

$$y = [8.78 - 6.67e^{-8.06t} - 3.09e^{-5.72t}\sin(38.5t + 0.755)]\cdot 10^{-6}. \tag{35}$$

The exact solution of initial equation (30) is given by [1]

$$y = [8.78 - 6.85e^{-8.3t} + 0.065e^{-65.9t} - 2.53e^{-4.42t}\sin(38.5t + 0.913)]\cdot 10^{-6}. \tag{36}$$

The curves for the exact and the approximate solutions are shown on Fig. 2. As is clear, the deviations between them are very significant.

Appendix III

In conclusion, we consider the lowering of the order of a linear differential equation in general form. We assume that we are given the equation

$$y''' + c_2y'' + c_1y' + c_0y = 0 \tag{37}$$

and the initial conditions (for t = 0) corresponding to it

$$y = 0, \quad y' = y_0' \quad y'' = 0\,.$$

We will assume that the system is stable, i.e., that the following conditions hold

$$c_0 > 0, \; c_1 > 0, \; c_2 > 0, \; c_1c_2 > c_0. \tag{38}$$

We now lower the order of equation (37). For this, by means of system (18) of equations and equation (20), we compute the integrals:

$$\int_0^\infty y\,dt = \frac{c_2y_0'}{c_0}\cdot \qquad \int_0^\infty y^2 dt = \frac{c_2^2 + c_0}{2c_0(c_1c_2 - c_0)}(y_0')^2,$$
$$\int_0^\infty (y')^2 dt = \frac{c_2^2 + c_1}{2(c_1c_2 - c_0)}(y_0')^2. \tag{39}$$

Then, by using system (17) of equations, we find the coefficients of the second-order approximation equation:

$$b_0 = \frac{c_0}{c_2}, \qquad b_1 = \frac{c_1 c_2 - c_0}{c_0 + c_2^2}. \tag{40}$$

As the result, the approximating equation takes the form:

$$y'' + \frac{c_1 c_2 - c_0}{(c_1 + c_2^2)} y' + \frac{c_0}{c_2} = 0, \tag{41}$$

where, for t = 0,

$$y = 0, \quad y' = y_0'.$$

Received March 16, 1959

LITERATURE CITED

[1] V. V. Solodovnikov, Yu. I. Topcheev and G. V. Krutikova, Frequency Methods of Constructing Transient Responses [in Russian] GITTL (1955).

[2] É. A. Meerovich, "Computing transient responses in complex electrical circuits," [in Russian] Izv. AN SSSR, otd. tekhn. n. number 10 (1950).

[3] É. A. Meerovich and V. A. Taft, "Engineering methods of computing nonsteady-state processes in complex energy systems," [in Russian], Électrichestvo 8 (1952).

[4] G. E. Pukhov, Complex Calculus [in Russian] (1956).

ON DESIGNING CORRECTING CIRCUITS FOR AUTOMATIC CONTROL SYSTEMS IN ACCORDANCE WITH THE MEAN SQUARE ERROR CRITERION*

V. I. Kukhtenko

(Moscow)

A method is presented for finding the optimal (in the sense of minimum mean square error) transfer function, the degree of whose numerator is a given number of units less than the degree of the denominator, and a discussion is given of the relationship of this question with the question of physical realizability of the correction circuit.

An example is provided.

The problem of designing control systems which are optimal in some sense is divided into two parts. First, it is necessary to determine the transfer function which is optimal from the point of view of the criterion chosen. Second, a method should be developed for synthesizing the optimal transfer function obtained by means of available actual elements. These two parts are different in their physical and mathematical content and are most frequently solved in isolation from one another.

In the makeup of automatic control systems there frequently enter elements whose presence is determined by requirements of design execution, character of operation, form of the energy sources provided, questions of reliability, etc., i.e., by conditions which relegate to the background of the design the dynamic properties of these elements. The dynamic quality of these "invariable" elements may turn out to be incompatible with the dynamic qualities of the optimal transfer function, defined by some criterion, in the sense that, for implementing the system as a whole, there might be required a physically nonrealizable correcting device.

In this paper we consider a method of obtaining an optimal transfer function in accordance with the criterion considered in [1] and essentially identical with the criterion of Zadeh and Ragazzini [2].

A control system as a whole, or its individual elements, are considered below to be physically realizable in those cases wherein the transfer functions corresponding to them have the form of rational fractions in which the degree of the numerator is less than, or equal to, the degree of the denominator. This limitation is imposed in addition to the ordinarily adopted condition of physical realizability given by the equality

$$k(t) = 0 \text{ for } t < 0, \tag{1}$$

where k(t) is the impulsive response of the system or its elements. The holding of equality (1) alone does not bespeak the possibility of realizing the system or the element. For example, the impulsive response of the ideal differentiating link, $k(t) = \delta(t)$, satisfies (1), nevertheless, in view of the unavoidable coordinate limitations and the inertia of all real devices, such a link in its pure form cannot be implemented.

---

* Presented at the All-Moscow Seminar on probabilistic methods of automatic control, held at the IAT AN SSSR on April 24, 1958.

We now consider the influence of the structure of the "invariable" portion of the system on the form of the transfer function of the closed system, bearing in mind that, as correcting links, we shall use physically realizable (in the sense given above) elements.

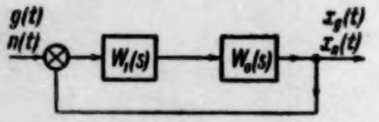


Fig. 1.

Figures 1 and 2 show two of the most frequently appearing schemes for correcting automatic control systems; on Fig. 1, by means of a correcting link in series and, on Fig. 2, by means of a feedback circuit. The transfer function of the object to be controlled (the "invariable" portion of the system) is denoted by $W_0(s)$, the transfer function of the correcting circuit by $W_1(s)$:

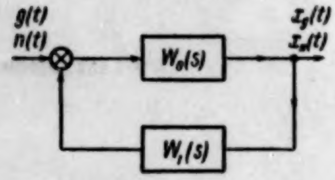


Fig. 2.

$$W_0(s) = \frac{N_0(s)}{M_0(s)}, \quad W_1(s) = \frac{P_1(s)}{H_1(s)}, \tag{2}$$

$$\begin{gathered} N_0(s) = a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0, \\ M_0(s) = b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0, \\ P_1(s) = c_p s^p + c_{p-1} s^{p-1} + \ldots + c_1 s + c_0, \\ H_1(s) = d_h s^h + d_{h-1} s^{h-1} + \ldots + d_1 s + d_0 \\ (n \leqslant m, \ p \leqslant h); \end{gathered} \tag{3}$$

We denote by $K(s)$ the transfer function of the closed-loop control system, and present it in the form

$$K(s) = \frac{L(s)}{R(s)},$$

where $L(s)$ and $R(s)$ are polynomials in $\underline{s}$ of degrees, respectively, l and $\underline{r}$.

For the system with the correcting link in series,

$$W_1(s) = \frac{L(s)\, M_0(s)}{N_0(s)\,[R(s) - L(s)]}, \quad p = l + m, \ h = n + r. \tag{4}$$

In this case, the correction link is realizable when $h - p \geq 0$, or

$$r - l \geqslant m - n. \tag{5}$$

For the system with the correction link in parallel,

$$W_1(s) = \frac{R(s)}{L(s)} - \frac{M_0(s)}{N_0(s)}. \tag{6}$$

In the left member of (6) there are two rational fractions, whereby the second allows an integral part to be isolated

$$\frac{M_0(s)}{N_0(s)} = \frac{b_m}{a_n} s^{m-n} + \ldots + \frac{M_1(s)}{N_0(s)}, \tag{7}$$

where the degree of $M_1(s)$ equals $\underline{n}$.

In order that the function $W_1(s)$ be realizable, it is necessary that it not contain an integral part, which must be cancelled by the corresponding multinomial which is separated from the fraction $R(s)/L(s)$.

This gives the following condition for physical realizability in the case of feedback:

$$r - l = m - n. \tag{5a}$$

Consequently, the limitation imposed on the form of the optimal transfer function in these two very common cases consists of this, that the degree of its numerator must be (m − n) less than the degree of its denominator.

We now form the integral:

$$T_0 = \int_0^\infty [k(t)]^2\, dt.$$

By the definition of the impulsive response,

$$\int_0^\infty \delta(t-\tau)\,\delta(t)\,dt = \delta(\tau) \quad \text{for} \quad \tau \geqslant 0$$

and

$$\int_0^\infty \delta(t)\,\delta(t)\,dt = \infty.$$

If it be required that the integral T_0 be finite, the function k(t) will then not have in its makeup a pulse constituent, the degree of whose numerator is one less than the degree of its denominator. Since differentiation in the t domain corresponds to multiplication in the s domain, the finitude of the integrals T_0 and

$$T_1 = \int_0^\infty [k'(t)]^2\, dt$$

will mean that the degree of the numerator of the function sK(s) is one less than the degree of its denominator; correspondingly, the degree of the numerator of the function K(s) is two less than the degree of its denominator.

Analogous reasoning leads to the conclusion that, for the degree of the numerator of K(s) to be q less than the degree of its denominator, it is necessary and sufficient that the following inequalities hold:

$$\begin{gathered} T_0 = \int_0^\infty [k(t)]^2\, dt < \infty, \\ T_1 = \int_0^\infty [k'(t)]^2\, dt < \infty \\ \cdots\cdots\cdots\cdots\cdots \\ T_{q-1} = \int_0^\infty [k^{(q-1)}(t)]^2\, dt < \infty. \end{gathered} \tag{8}$$

For the circuits shown in Figs. 1 and 2, let n(t) be the noise, given as a stationary random function of time, and let it be known of the signal, g(t), that its Laplace transform, $g^*(s)$, is a rational fraction. We note that everything presented below is extended, by analogy, to the case when the signal contains a stationary random component, m(t).

We shall call an automatic control system optimal if [1]:

1) the mean square error

$$\overline{x_{n}^{2}} = M\,[x_{n}^{2}(t)] = \int_0^\infty\int_0^\infty R_n(\tau-\theta)\,k(\tau)\,k(\theta)\,d\tau\,d\theta \tag{9}$$

(where $R_n(t)$ is the correlation function of the noise and k(t) is the impulsive response of the closed-loop system) is minimal;

2) the instantaneous value of the error component corresponding to the forced motions [3] engendered by the signal g(t) equals zero:

$$\varepsilon_f(t) = g(t) - x_g \text{ for } (t) = 0. \tag{10}$$

Condition (10) means that, after termination of the transient response, the signal, g(t), passes through the system without distortion.

In accordance with [1], in order to satisfy (10) the impulsive response k(t) must satisfy the following equalities:*

$$K(s_i+\Delta_i) = \int_0^\infty k(t)\,e^{-(s_i+\Delta_i)t}\,dt = 1-\bar{\varepsilon}_{0i},$$

$$\left[\frac{\partial K(s)}{\partial s}\right]_{s=s_i+\Delta_i} = -\int_0^\infty t k(t)\,e^{-(s_i+\Delta_i)t}\,dt = \bar{\varepsilon}_{1i},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left[\frac{\partial^{m_i-1}K(s)}{\partial s^{m_i-1}}\right]_{s=s_i+\Delta_i} = (-1)^{m_i-1}\int_0^\infty t^{m_i-1}k(t)\,e^{-(s_i+\Delta_i)t}\,dt = \bar{\varepsilon}_{m_i-1,\,i}$$

$$(i=1,2,\ldots,\nu),$$

$$\operatorname{Re}\,[K(\alpha_k+i\beta_k+\Delta_k)] = \int_0^\infty k(t)\,e^{-(\alpha_k+\Delta_k)t}\cos\beta_k t\,dt = 1-\bar{\varepsilon}_{01k},$$

$$\operatorname{Im}\,[K(\alpha_k+i\beta_k+\Delta_k)] = -\int_0^\infty k(t)\,e^{-(\alpha_k+\Delta_k)t}\sin\beta_k t\,dt = \bar{\varepsilon}_{02k}, \tag{11}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\operatorname{Re}\left[\frac{\partial^{m_k-1}K(s)}{\partial s^{m_k-1}}\right]_{s=\alpha_k+i\beta_k+\Delta_k} = (-1)^{m_k-1}\int_0^\infty t^{m_k-1}k(t)\,e^{-(\alpha_k+\Delta_k)t}\times$$

$$\times\cos\beta_k t\,dt = \bar{\varepsilon}_{m_k-1,1k},$$

$$\operatorname{Im}\left[\frac{\partial^{m_k-1}K(s)}{\partial s^{m_k-1}}\right]_{s=\alpha_k+i\beta_k+\Delta_k} = (-1)^{m_k}\int_0^\infty t^{m_k-1}k(t)\,e^{-(\alpha_k+\Delta_k)t}\times$$

$$\times\sin\beta_k t\,dt = \bar{\varepsilon}_{m_k-1,2k} \quad (k=1,2,\ldots,\mu,\ \nu+\mu=n).$$

Here, Re denotes the real part, Im denotes the imaginary part, ν is the number of real poles, s_i, of the image $g^*(s)$, μ is the number of complex poles $s_k = \alpha_k \pm i\beta_k$ of the image $g^*(s)$, m_i and m_k are the multiplicities of the corresponding poles, Δ_i and Δ_k are arbitrary positive numbers, so chosen that the following condition holds

$$\operatorname{Re}(s_j+\Delta_j) > 0. \tag{12}$$

*The equation corresponding to distortion indicator $\bar{\epsilon}_{2k}$ was omitted in [1].

The quantities $\bar{\epsilon}_{\rho\xi}$, the so-called distortion indicators, determine the accuracy of transmission of the controlled signal $\bar{g}(t)$, distinguished from g(t) by the fact that its image, $\bar{g}^*(s)$, has poles $s_j + \Delta_j$. The meaning of equations (11) consists in this, that they establish some still undefined measure of distortion of the fictitious signal $\bar{g}(t)$ in its passage through the system K(s), which can then be so chosen that the basic signal, g(t), is reproduced in accordance with (10). For this, it is necessary that one set $\Delta_j = \bar{\epsilon}_{\rho\xi} = 0$ in (11).

As is obvious from the discussion above, the problem of determining the optimal impulsive response consists in minimizing the functional in (9), with additional limitations of the form of (8) and (11). Problems of this type were often considered in the theory of optimal systems [4, 5]. Investigation of the first variation of the functional

$$\bar{I}_0[k(t)] = \bar{x}_n^2 - 2\sum_{i=1}^{n}\sum_{j=0}^{m_i-1}\gamma_{ij}K^j(\bar{s}_i) + \sum_{i=0}^{q-1}\lambda_r\int_0^{\infty}[k^{(r)}(t)]^2\,dt, \tag{13}$$

where $K^j(s_i)$ conventionally denotes the right member of (11), leads to the following equation, of the Wiener-Hopf type, for determining $k(t)$:

$$\begin{gathered}\int_0^{\infty}k(\tau)R_\sigma(t-\tau)\,d\tau - \sum_{i=1}^{n}\sum_{j=0}^{m_i-1}(-1)^j t^j e^{-(\alpha_i+\Delta_i)t}\times \\ \times[\gamma_{ij}\cos\beta_i t - \bar{\gamma}_{ij}\sin\beta_i t] = 0 \quad \text{for} \quad t>0,\end{gathered} \tag{14}$$

$$R_\sigma(t) = R_n(t) + \sum_{r=0}^{q-1}\lambda_r\delta^{2r}(t). \tag{15}$$

Here, λ_r and γ_{ij} are Lagrange multipliers.

The solution of equation (14) is carried out by the ordinary methods [5, 6].

Application of Parseval's formula to the second variation gives a sufficient condition for a minimum

$$S_\sigma(\omega) \geqslant 0, \tag{16}$$

where $S_\sigma(\omega)$ is the spectral density corresponding to the correlation function $R_\sigma(t)$.

The expression for the mean value of the square of the noise of the output provided by an optimal system, has the form:

$$\bar{x}_n^2 = \sum_{i=1}^{n}\sum_{j=0}^{m_i-1}\gamma_{ij}K^j(\bar{s}_i) - \sum_{r=0}^{q-1}(-1)^r\lambda_r\int_0^{\infty}k(t)\,k^r(t)\,dt. \tag{17}$$

The Lagrange multipliers, γ_{ij}, from the expression for the optimal transfer function are excluded by equations (11), in which it is necessary to set $\Delta_j = \bar{\epsilon}_{\rho\xi} = 0$.

In the general case, decomposition of the function $S_\sigma(\omega)$ into two complex conjugate factors is impossible since, for this, one would have to solve an algebraic equation of high degree with unknown coefficients, which is impossible. Therefore, instead of λ_r, it is necessary to introduce new unknown coefficients, b_r, by the formula

$$\begin{gathered}P(\omega) = a_{2l}(\lambda_r)\,\omega^{2l} + a_{2(l-1)}(\lambda_r)\,\omega^{2(l-1)} + \ldots \\ \ldots + a_2(\lambda_r)\,\omega^2 + a_0(\lambda_r) = a_{2l}(\lambda_r)\prod_l(b_l + i\omega)(b_l - i\omega),\end{gathered} \tag{18}$$

where $P(\omega)$ is the numerator of the rational-fraction function $S_\sigma(\omega)$.

After determining K(s) by formulae (4) and (6), one can find the transfer function, $W_1(s)$, of the correcting circuit, which will depend on the undetermined parameters b_l. The values of these parameters should now be so chosen that the coefficients of the integral part in the expression for $W_1(s)$ become equal to zero (with the exception of the coefficient of s^0). Then, the transfer function $W_1(s)$ will have a numerator whose degree is not greater than that of the denominator, and, consequently, will satisfy the criterion of physical realizability adopted above.

Example. Let the executive element in a control system be an electric motor with the transfer function

$$W_0(s) = \frac{K}{s}, \quad K = 3\,\frac{1}{\text{sec}\cdot\text{volt}}.$$

The signal has the form:

$$g(t) = g_0 + g_1 t,$$

and the correlation function of the noise is

$$R_n(t) = R_0 e^{-a|t|}, \quad R_0 = 1, \quad a = 1\,\frac{1}{\text{second}}.$$

In accordance with the general theory for obtaining physically realizable correcting circuits, the degree of the numerator of the optimal transfer function must, in this case, be one less than the degree of the denominator, which leads to the equation

$$(\gamma_0 - \gamma_1 t)\, e^{\Delta t} - \int_0^\infty R_\sigma(t-\tau)\, k(\tau)\, d\tau = 0,$$

where $R_\sigma(t) = R_0 e^{-a|t|} + \lambda_0 \delta(t)$.

If we use, for the solution, the formula

$$K(i\omega) = \frac{1}{\Psi(i\omega)}\left[\frac{B^*(i\omega)}{\Psi(-i\omega)}\right]_+$$

and take into account that

$$B^*(i\omega) = \frac{\gamma_0}{i\omega + \Delta} - \frac{\gamma_1}{(i\omega + \Delta)^2}, \qquad S_\sigma(\omega) = \lambda_0\,\frac{(b + i\omega)(b - i\omega)}{(a + i\omega)(a - i\omega)},$$

we get

$$K(s) = \frac{As^2 + (Aa + E)\, s + Ea}{s^3 + (2\Delta + b)\, s^2 + \Delta\,(2b + \Delta)\, s + \Delta^2 b}.$$

The conditions for undistorted reproduction of the signal g(t) become transformed to the conditions for obtaining a system with second order astatism, which gives

$$K(s) = \frac{As^2 + \Delta\,(2b + \Delta)\, s + \Delta^2 b}{s^3\,(2\Delta + b)\, s^2 + \Delta\,(2b + \Delta)\, s + \Delta^2 b},$$

$$A = \frac{[\Delta\,(a - b)\,\Delta + 2ab]}{a^2}.$$

The transfer function of the feedback circuit (Fig. 2) is

$$W_1(s) = \frac{1}{K(s)} - \frac{s}{K} = \frac{1}{A}s + \frac{\left[(2\Delta + b) - \frac{\Delta}{A}(2b+\Delta)\right]s^2 + \left[(2b+\Delta)\Delta - \frac{\Delta^2}{A}b\right]s + \Delta^2 b}{As^2 + (2b+\Delta)\Delta s + \Delta^2 b} - $$

$$- \frac{s}{K}.$$

The conditions of physical realizability have the form:

$$A = K,$$

$$W_1(s) = \frac{\left[(2\Delta + b) - \frac{\Delta}{A}(2b + \Delta)\right]s^2 + \left[(2b + \Delta)\Delta - \frac{\Delta^2}{A}b\right]s + \Delta^2 b}{As^2 - (2b + \Delta)\Delta s + \Delta^2 b}$$

$$b = \frac{a}{\Delta}\,\frac{Ka - \Delta^2}{2a - \Delta}.$$

For choosing the value of Δ it is necessary to take three factors into consideration: the duration of the transient response, the satisfying of inequality (16) and the magnitude of $\overline{x_n^2}$ [1].

Computations provide the formula for the mean-square error

$$\overline{x_n^2} = \frac{A^2a^2B + A^2aC - 2aACD - 2ACa^2 + Db^2a + B^2a^2 - C^2 + D^2aC + Da^2C + DBC}{(aB + C + Da^2 + a^3)(DB - C)}$$

$$B = \Delta(2b + \Delta), \quad C = \Delta^2 b, \quad D = 2\Delta + b.$$

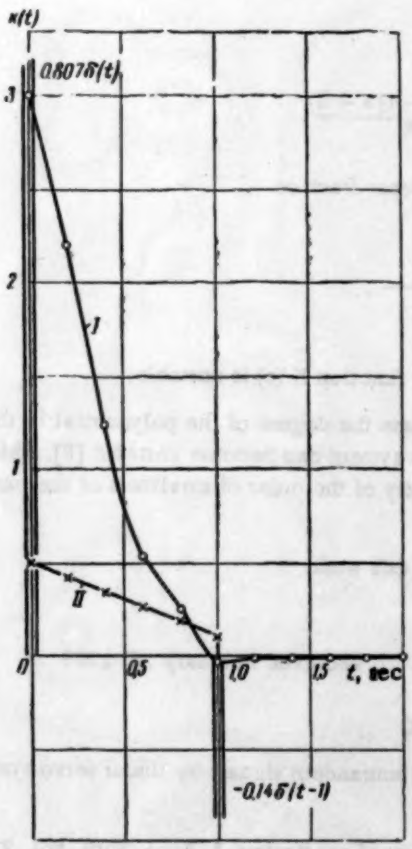


Fig. 3.

By setting $\Delta = 3$, we find that $b = 2$. From the expression for the optimal transfer function it is obvious that the duration of the transient response in this case is of the order of one or two seconds.

A sufficient condition for a minimum has the form: $\lambda_0 > 0$, which leads to the inequality $b > a$ which, for $a = 1$ and $b = 2$, is satisfied.

The final expressions for the complete transfer function and for the transfer function of the feedback path will be

$$K(s) = \frac{3s^2 + 21s + 18}{s^3 + 8s^2 + 21s + 18},$$

$$W_1(s) = \frac{s^3 + 15s + 18}{3s^2 + 21s + 18}.$$

The transfer function of the feedback link can be realized by means of a passive RC-circuit [7].

For the parameter values chosen, the mean value of the squared noise at the output is

$$\overline{x_n^2} \approx 1{,}02.$$

The impulsive response corresponding to transfer function K(s) is written in the form

$$K(t) = \frac{A-b-\Delta}{b-\Delta}\Delta^2 te^{-\Delta t} + \frac{A\Delta^2-2Ab\Delta+2\Delta b^2}{(b-\Delta)^2}e^{-\Delta t} + \frac{b^2(A-2\Delta)}{(\Delta-b)^2}e^{-bt} = 18te^{-3t} + 15e^{-3t} - 12e^{-2t}.$$

The graph of K(t) is shown on Fig. 3 (curve I). For purposes of comparison, Fig. 3 also shows the graph of the optimal impulsive response (II) as determined by the method of Zadeh-Ragazzini [2], giving, for the same conditions, a value $\overline{\varepsilon^2_{nz}}$ which is approximately equal to $\overline{\varepsilon^2_{n}}$. On the basis of [5], for T = 1, $\overline{\varepsilon^2_{nz}} \approx 0.984$ and $k_z(t) = 0.492 - 0.317t + 0.8\,\delta(t) - 0.14\,\delta(t-1)$. It is clear from Fig. 3 that the durations of the transient responses for k(t) and $k_z(t)$ are approximately identical.

It should be mentioned that the attempt to obtain a realizable transfer function of the correcting link approximately on the basis of an optimal transfer function which does not possess property (5) leads to significant difficulties. For example, for the circuit of Fig. 2, let the optimal transfer function have the form:

$$K(s) = \frac{As^2 + Bs + 1}{Cs^2 + Ds + 1}.$$

Then, the transfer function of the correcting circuit is written as follows:

$$W_1(s) = \frac{-As^3 + (KC - B)s^2 + (KD - 1)s + K}{(As^2 + Bs + 1)K}.$$

This transfer function is unrealizable. If it is approximated to by a realizable function in the natural way, i.e., if we introduce into its denominator the term xs^3 with a small coefficient $\underline{x}$, then the new transfer function, $W_1'(s)$ becomes equal to the improper fraction

$$W_1'(s) = \frac{-As^3 + (KC - B)s^2 + (KD - 1)s + K}{xs^3 + KAs^2 + KBs + K},$$

and the transfer function of the entire system becomes equal to the proper fraction

$$K'(s) = \frac{\frac{x}{K}s^3 + As^2 + Bs + 1}{\frac{x}{K^2}s^4 + Cs^2 + Ds + 1}.$$

It is clear from this last expression that the system with transfer function K'(s) is unstable.

Generally, with the introduction of small parameters with increase the degree of the polynomial in the denominator of the transfer function of the correcting circuit, a stable system can become unstable [8], which makes the method of natural approximation inapplicable, independently of the order of smallness of the parameters introduced.

In conclusion, I wish to thank V. S. Pugachev for his interest in this work.

Received February 26, 1959

LITERATURE CITED

[1] V. I. Kukhtenko, "On the joint reproduction of random and nonrandom signals by linear servo systems," [in Russian] Automation and Remote Control (USSR) 17, 5 (1956).*

[2] L. A. Zadeh and J. R. Ragazzini, Extension of Wiener's theory of prediction, J. Appl. Phys. vol. 21, No. 7, 1950.

* See English translation.

[3] S. P. Strelkov, "Concerning the general theory of linear amplifiers, part 1," [in Russian] Automation and Remote Control (USSR) 9, 3 (1948).

[4] V. S. Pugachev, The Theory of Random Functions and their Application to Automatic Control Problems, [in Russian] (Gostekhizdat (1959).

[5] V. V. Solodovnikov, Introduction to the Statistical Dynamics of Automatic Control Systems [in Russian] Gostekhizdat (1952).

[6] Syue-Sen' Tsyan', Engineering Cybernetics [Russian Translation] ILI (1956).

[7] Automatic Control Fundamentals, Vol. 2, Part 2 [in Russian] edited by V. V. Solodovnikov. Mashgiz (1959).

[8] M. V. Meerov, Introduction to the Dynamics of Automatic Control of Electrical Machines [in Russian] Izdatelstvo AN SSSR (1956).

THE EFFECT OF NOISE ON PHASE-LOCKED OSCILLATOR OPERATION

V. I. Tikhonov

(Moscow)

The action of external noise and of inherent fluctuations on an inertialess phase-locked oscillator are considered by means of the Fokker-Planck equation. Formulae for the mean and dispersion of the synchronized generator's frequency are found. The analogy is established between the problem considered here and the problem of synchronizing a self-excited oscillator in the presence of noise.

Today, in various devices of radio engineering and automation, widespread usage is made of phase-locked oscillators for automatic frequency control (A.F.C.), the theory of operation of which can be elucidated using the example of the simplest functional schematic (Fig. 1).

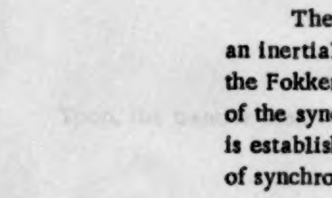


Fig. 1. A.F.C. functional schematic.

The harmonic oscillations, $u_f(t)$ of the fundamental generator (FG) and $u_s(t)$ of the synchronized generator (SG), act on phase detector PD at whose output there appears a voltage which depends on the difference of phase of the sinusoids $u_f(t)$ and $u_s(t)$. This voltage, by means of low-frequency filter LFF and reactance modulator RM, changes the frequency of the synchronized generator, causing it to coincide with the frequency of the fundamental generator. With this, there is established a constant difference of phase between the sinusoids of the fundamental and the synchronized generators, which provides synchronized operation of both generators [1].

The operation of an A.F.C. circuit is considerably complicated if the presence of noise is taken into account. Under practical conditions of A.F.C. operation there occur both external noise and fluctuations which are organically inherent in the circuit elements themselves. As is well-known [2-4], anode current noise in the generator tubes, random variations in the voltage supply and other causes result in fluctuations of the generators' frequency (phase). Moreover, together with the sinusoid from the fundamental generator, additional noise, n(t), may act on the phase detector, particularly if the fundamental generator is spatially remote from the rest of the A.F.C. elements.

The problem of the effect of noise on A.F.C. operation is, in general, very complicated mathematically because of the specific nonlinearity of the A.F.C. differential equation. Although several works [5-8] have been devoted to analyses of particular cases of the action of noise on A.F.C. systems, it is impossible to consider, even from a qualitative point of view, that the problem has been solved with sufficient completeness.

In the sequel, we shall limit ourselves to the consideration of a simplified, inertialess A.F.C. system for which the block schematic of Fig. 1 applies with the differences that the PD is replaced by a multiplying device and the LFF is considered to be an ideal low-frequency filter (its transfer constant equalling unity for low frequencies and zero for high frequencies). With this, we shall take into account the fluctuations of frequency of the fundamental and synchronized generators and the presence of the noise, n(t). This noise is assumed to be normal with zero mean; the spectral density of this noise is symmetric with respect to the mean frequency of the fundamental generator.

Posing of the Problem

If we ignore amplitude fluctuations, we can write the oscillations of the fundamental and synchronized generators in the following forms, respectively:

$$u_f(t) = A_1 \cos \Phi_1(t) = A_1 \cos[\omega_f t + \theta_1(t)], \tag{1}$$

$$u_s(t) = A_2 \sin \Phi_2(t) = A_2 \sin[\omega t + \theta_2(t)], \tag{2}$$

where A_1 and A_2 are constant amplitudes, ω_0 and ω are the mean frequencies and $\theta_1(t)$ and $\theta_2(t)$ are the random phases of the generators' oscillations.

The normal fluctuating noise n(t), with spectral density symmetric about frequency ω_0, can be presented in the form

$$n(t) = E(t) \cos \Phi(t) = E(t) \cos[\omega f t + \theta(t)], \tag{3}$$

where E(t) is the envelope of the fluctuations and $\theta(t)$ is the random phase.

The total voltage of the fundamental generator and the noise which acts on the multiplier block can be written as

$$u_f(t) + n(t) = A_1 \cos \Phi_1(t) + E_s(t) \cos \Phi_1(t) - E_{\mathrm{L}}(t) \sin \Phi_1(t),$$

where

$$E_s(t) = E(t) \cos(\theta - \theta_1), \quad E_t(t) = E(t) \sin(\theta - \theta_1). \tag{4}$$

By using the same methodology as in [9], we may easily show that the sinusoid $E_t(t)$ and the cosinusoid $E_s(t)$ composing the envelope E(t) have normal probability densities with zero means. If we denote the correlation function of the noise n(t) by

$$k_n(\tau) = \sigma^2 r(\tau) \cos \omega_0 \tau,$$

where σ^2 is the dispersion of the noise n(t), then the auto-correlation and cross-correlation functions of E_t and E_s are defined by the relationships

$$\begin{aligned} k(\tau) &= \overline{E_t(t) E_t(t+\tau)} = \overline{E_s(t) E_s(t+\tau)} = \sigma^2 r(\tau), \\ \lambda(\tau) &= \overline{E_s(t) E_s(t+\tau)} = -\overline{E_s(t) E_s(t+\tau)}. \end{aligned} \tag{5}$$

We denote the multiplier's transformation coefficient by μ. Then, the voltage at the multiplier's output will be

$$\begin{aligned} u_m(t) = \mu[u_f(t) + n(t)]\, u_s(t) = \frac{1}{2} \mu A_2 \{(A_1 + E_s)[\sin(\Phi_2 - \Phi_1) + \\ + \sin(\Phi_1 + \Phi_2)] - E_t[\cos(\Phi_2 - \Phi_1) - \cos(\Phi_1 + \Phi_2)]\}. \end{aligned} \tag{6}$$

Since the LFF is ideal, i.e., its transfer constant is unity for low frequencies $(\omega - \omega_0)$ and zero for high frequencies $(\omega + \omega_0)$, the voltage at the input to the RM will be

$$u(t) = \frac{1}{2} \mu A_2 [(A_1 + E_s) \sin \varphi - E_t \cos \varphi], \quad \varphi = \Phi_2 - \Phi_1. \tag{7}$$

Let the average frequency ω of the synchronized generator depend linearly on the voltage at the reactance modulator's input, i.e.,

$$\omega = \omega_{s0} - s u(t), \tag{8}$$

where ω_{a0} is the average frequency of the synchronized generator for u(t) = 0 and s is the slope of the linear portion of the reactance modulator's modulation characteristic.

From expressions (1), (2) and (7) we obtain

$$\frac{d\varphi}{dt} = \dot{\varphi} = \dot{\Phi}_2 - \dot{\Phi}_1 = (\omega_{a0} - \omega_0) - su(t) + (\dot{\theta}_2 - \dot{\theta}_1)$$

or

$$\dot{\varphi} = \Delta_0 - \Delta \sin\varphi - \frac{1}{A_1}\Delta(E_s \sin\varphi - E_c \cos\varphi) + \dot{\psi}, \tag{9}$$

where Δ_0 is the initial tuning of the generators,

$$\Delta_0 = \omega_{a0} - \omega_0, \quad \Delta = \frac{1}{2}\mu s A_1 A_2, \quad \dot{\psi} = \dot{\theta}_2 - \dot{\theta}_1. \tag{10}$$

To compute the statistical characteristics of $\dot{\varphi}$ from the nonlinear differential equation (9), in which the random functions $E_c(t)$, $E_s(t)$ and $\dot{\psi}(t)$ enter is, in the general case, quite difficult. By means of the substitution $x = \tan(\varphi/2)$, this equation can be reduced to the Ricatti equation and this can then be transformed into a linear homogeneous, second-order equation with variable coefficients. However, these transformations do not give any simplifications in principle.

In the sequel we shall apply to equation (9) the apparatus of Markoff processes, which is valid if the following inequality holds

$$\frac{2}{\mu s A_2 (A_1 + E_s)} \gg \tau_{kn} + \tau_{k\dot{\psi}}, \tag{11}$$

where τ_{kn} and $\tau_{k\dot{\psi}}$ are the correlation times on n(t) and $\dot{\psi}(t)$ respectively. This inequality holds in many practical cases, for example, for radio communication equipment.

Solution of the Fokker-Planck Equation

If we have the first-order differential equation

$$\dot{\varphi} = F[\varphi, \xi(t)], \tag{12}$$

where $\xi(t)$ is a stationary random function then, given a condition analogous to (11), the following Fokker-Planck equation (10) corresponds to it, defining the one-dimensional probability density for the random function $\varphi(t)$:

$$\dot{w}(\varphi) = -\frac{\partial}{\partial\varphi}[K_1(\varphi)w(\varphi)] + \frac{1}{2}\frac{\partial^2}{\partial\varphi^2}[K_2(\varphi)w(\varphi)]. \tag{13}$$

The structural numbers, K_1 and K_2, are computed from the formulae

$$K_1(\varphi) = \overline{F[\varphi, \xi(t)]}, \quad K_2(\varphi) = \int_{-\infty}^{\infty} \{\overline{F[\varphi, \xi(t)]F[\varphi, \xi(t+\tau)]} - K_1^2(\varphi)\}\, d\tau. \tag{14}$$

For the statistical averaging, denoted by the superscript bar, φ is considered as a fixed parameter, not as a random variable.

We now introduce a physically graphical interpretation of the Fokker-Planck equation (13) which we shall use in the sequel.

In each concrete realization of the process described by equation (12), the trajectory $\varphi(t)$ can have a very complicated form. The representative point moves along the t axis like a Brownian particle (i.e., like a particle engaged in Brownian motion). If we take a large number of realizations of the random process, we shall have a large number of representative points carrying out random wanderings. The points form, as it were, a "gas" undergoing diffusion, where the density of this "gas" at any point is proportional to the probability density. Each individual representative point constitutes a "molecule" of this "gas." It moves, but cannot be created or annihilated.

In correspondence with equation (13), the average velocity of a "molecule" at a chosen point of the φ axis equals $K_1(\varphi)$. By multiplying this by the relative density of the "molecules," $w(\varphi)$,, we obtain the convection flux of the "molecules," equal to $K_1(\varphi)w(\varphi)$. The presence of the concentration gradient, $\partial w/\partial \varphi$, also leads to the appearance of a diffusion flux, which is proportional to the concentration gradient. The total flux of "molecules," or the probability across section φ in unit time, equals

$$G(\varphi) = K_1(\varphi)\,w(\varphi) - \frac{1}{2}\frac{\partial}{\partial\varphi}\left[K_2(\varphi)\,w(\varphi)\right]. \tag{15}$$

Thus, Equation (13) can be written in the form

$$\dot{w}(\varphi) + \frac{\partial}{\partial\varphi}G(\varphi) = 0 \tag{16}$$

and may be considered as the equation of probability conservation.

In the sequel we shall only be interested in stationary distributions, characterized by the fact that the probability distribution, $w(\varphi)$, does not depend on time, i.e., $\dot{w}(\varphi) = 0$. Therefore, we obtain for it, from (13), the following equation

$$\frac{1}{2}\frac{d^2}{d\varphi^2}\left[K_2(\varphi)\,w(\varphi)\right] - \frac{d}{d\varphi}\left[K_1(\varphi)\,w(\varphi)\right] = 0. \tag{17}$$

By using formulae (5) and (14) for computing K_2, one can easily show that, for equation (9), the structural numbers equal

$$K_1(\varphi) = \Delta_0 - \Delta\sin\varphi, \quad K_2(\varphi) = \int_{-\infty}^{\infty}\left[\left(\frac{\Delta}{A_1}\right)^2 k(\tau) + k_{\dot{\psi}}(\tau)\right]d\tau. \tag{18}$$

Thus, as applied to the problem under consideration, equation (17) can be written in the form

$$\frac{\partial^2 w}{\partial\varphi^2} - \frac{d}{d\varphi}\left[(D_0 - D\sin\varphi)\,w\right] = 0, \tag{19}$$

where

$$D_0 = \frac{2\Delta_0}{K_2} = \frac{2(\omega_{20} - \omega_0)}{K_2}, \quad D = \frac{2\Delta}{K_2} = \frac{\mu s A_1 A_2}{K_2}. \tag{20}$$

It should be mentioned that, in the given case, K_2 does not depend on φ, the fluctuations of the generator frequencies and the noise entering additively into the fundamental signal channel, the sole difference being that the noise is multiplied by a factor which takes into account the translation of the noise from the multiplier input to the input of the synchronized generator. In view of this, we obtain a complete analogy of the problem under consideration with the problem of synchronizing a self-excited oscillator in the presence of noise [11].

The general solution of equation (19) can be written in the form

$$w(\varphi) = C_1 e^{D_0\varphi + D\cos\varphi}\int_{C_2}^{\varphi} e^{-D_0\psi - D\cos\psi}\,d\psi.$$

It contains two constants of integration, C_1 and C_2, which are defined by two conditions: 1) the condition of periodicity, $w(\varphi \pm 2\pi) = w(\varphi)$; 2) the normalization condition for the probability density in each period,

$$\int_0^{2\pi} w(\varphi)\, d\varphi = 1.$$

By carrying out the necessary computations we get [11]

$$w(\varphi) = \frac{1}{N} e^{D_0\varphi + D\cos\varphi} \int_{\varphi}^{\varphi+2\pi} e^{-D_0\psi - D\cos\psi}\, d\psi, \tag{21}$$

$$N = 4\pi^2 e^{-\pi D_0} |I_{iD_0}(D)|^2. \tag{22}$$

Here $I_{i\nu}(z)$ is a Bessel function of imaginary order and imaginary argument.

With zero initial tuning, $D_0 = 0$, the probability density for phase is determined by the expression

$$w(\varphi) = \frac{1}{2\pi I_0(D)} e^{D\cos\varphi}$$

and is symmetric with respect to the values $\varphi = \pm 2k\pi$ ($k = 0, 1, 2, \ldots$), corresponding to stable operation of the A.F.C. and the absence of noise. For $D_0 \neq 0$, the form of the curve for $w(\varphi)$ is asymmetric. It can be found by using tables of Bessel functions of imaginary order and imaginary argument [12] and carrying out the numerical integration of (21).

Statistical Characteristics

For the qualitative behavior of the phase difference, φ, in the presence of noise, one can give the following graphic explanation, suggested by R. L. Stratonovich. We first assume that the noise, $n(t)$, is absent, i.e., $E_t = E_s = 0$. Then, equation (9) takes the form:

$$\dot{\varphi} = \Delta_0 - \Delta\sin\varphi + \dot{\psi}. \tag{23}$$

We shall formally consider the quantity $s = \int_0^t \varphi(x)dx$ as the path followed by some point during time t

and correspondingly, $\dot{\psi}$ will be considered the point's acceleration. Then, equation (23) can be interpreted as the equation of motion of a point in a field of regular external forces which depend nonlinearly on position, with the presence of the random disturbance $\dot{\psi}$. If $\Delta_0 - \Delta\sin\varphi$ is considered as the strength of the external force field, then the potential of these forces equals

$$U = -\int_0^{\varphi} (\Delta_0 - \Delta\sin\varphi)\, d\varphi = -\Delta_0\varphi - \Delta\cos\varphi.$$

This potential has the same form as the potential of a mass point moving on an undulating sloping surface (Fig. 2). This surface, for $\Delta_0 > 0$ and $\Delta > \Delta_0$, has "pockets" in the neighborhoods of the points $\varphi = \pm 2k\pi$ ($k = 0, 1, 2, \ldots$), these being stable positions of the point in the absence of random disturbances. At some initial moment of time, let the point be at "pocket" M_1. A random disturbance, first of all, moves the point in the neighborhood of the "pocket" in a random fashion, causing small fluctuations of frequency and phase, and, secondly, more or less frequently takes it from the "pocket," after which it gradually rolls to the following next lower "pocket," M_2. Less frequently, a random shock will throw the point into "pocket" M_3 which is placed at a higher level. The first transition corresponds to an increase of the phase by 2π, the second to a decrease by the same amount. In

general, there is a sliding off of points from the sloping surface which occurs the more rapidly, the larger is the slope (the initial tuning), the shallower are the "pockets" (the lower is the amplitude of the synchronizing voltage) and the larger are the random disturbances. This means that there occurs a regular change of the synchronized generator's average frequency, causing it to approximate to the frequency of the fundamental generator.

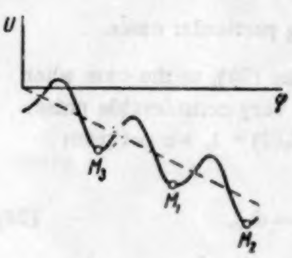


Fig. 2. Potential contour.

If the random disturbance, $\dot\psi$, is very small, then the transitions from M_1 to M_2 and M_3 will have low probability and, in practice, need not be taken into account. With this, the magnitude of φ will experience small fluctuations in the neighborhood of stable state M_1. For small disturbances, independently of their structure, and if condition (11) holds, one can limit oneself to considering, instead of equation (9), a linearized equation with respect to the fluctuating quantities [13-15].

For $\Delta < \Delta_0$, the "pockets" vanish, due to the overly large slope and, with this, the point uninteruptedly slides down, and a synchronized mode of operation is impossible.

If one takes into account the presence of noise in the channel of the fundamental generator ($E_t \neq 0$, $E_s \neq 0$), then the slope of the undulating surface will experience fluctuations of the standing-wave type, with random amplitudes distributed normally with zero mean. If condition (11) holds, these random pulsations occur very rapidly. They also make possible a systematic translation of the point to the average slope of the surface.

We note that the idea of jumps of the representative points in systems similar to the ones we are considering, is contained in [16].

We now compute the mean frequency of the synchronized generator. From equation (9) we have

$$\bar{\dot\varphi} = \overline{(\Delta_0 - \Delta \sin\varphi)} - \frac{\Delta}{A_1}\overline{(E_s \sin\varphi - E_t \cos\varphi)} + \bar{\dot\psi}.$$

According to condition (11), the random functions E(t) and $\dot\psi(t)$ vary rapidly in time in comparison with $\varphi(t)$. Therefore, despite the correlation between E(t), $\dot\psi(t)$ and $\varphi(t)$ which, in theory, can be taken into account, we shall, with an admissible approximation, first average the rapidly varying quantities separately, and then the slowly varying ones. We note, by the way, that the same approach can also be used in computing the dispersion of the frequency.

By taking into account that $E_t = E_s = 0$ and $\bar{\dot\psi} = 0$, we can write

$$\bar{\dot\varphi} = \overline{(\Delta_0 - \Delta \sin\varphi)} = \int_0^{2\pi} (\Delta_0 - \Delta \sin\varphi)\, w(\varphi)\, d\varphi. \tag{24}$$

It follows from formulae (15) and (18) that

$$(\Delta_0 - \Delta \sin\varphi)\, w(\varphi) = K_1(\varphi)\, w(\varphi) = G + \frac{1}{2} K_2 \frac{dw(\varphi)}{d\varphi}. \tag{25}$$

If we substitute (25) in (24) and bear in mind that, in the stationary state, the flux G is constant but the probability density is a periodic function with period 2π, we get that

$$\bar{\dot\varphi} = 2\pi G. \tag{26}$$

If we use expression (21) for computing the flux with (15), we obtain the final formula

$$\bar{\dot\varphi} = \bar\omega - \omega_0 = \frac{\pi K_2}{N}\left(1 - e^{-2\pi D_0}\right) = \Delta_0 \frac{\operatorname{sh} \pi D_0}{\pi D_0} \left| I_{iD_0}(D) \right|^{-2}. \tag{27}$$

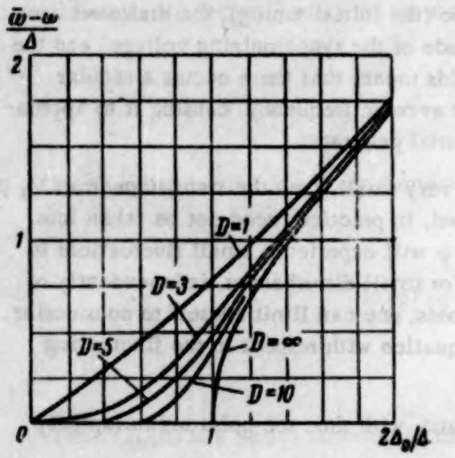


Fig. 3. Difference of the average frequencies as a function of the circuit parameters and the characteristics of the noise.

The results of computations carried out by means of this formula by I. G. Akopyan are shown on Fig. 3.

We now consider two limiting particular cases.

1. As is obvious from formulae (20), to the case when $D \to 0$ and $D_0 \to 0$ there corresponds very considerable noise. By using the well-known equality, $I_0(0) = 1$, we get from formula (27) that

$$\overline{\omega} = \omega_f + \Delta_0 = \omega_{a0}. \tag{28}$$

Thus, in the case of a very high noise level, the synchronizing signal has no effect, and the average frequency of the synchronized generator remains equal to its initial value ω_{a0}.

2. It can be shown that in the opposite limiting case, when the noise level is very low, the following dependency, obtained by Yu. B. Kobzarev and R. V. Khokhlov [17], is valid:

$$\overline{\omega} - \omega_f = \begin{cases} 0 & \text{for } \Delta_0 \leqslant \Delta, \\ (\Delta_0^2 - \Delta^2)^{1/2} & \text{for } \Delta_0 > \Delta. \end{cases} \tag{29}$$

It follows from this that, for a very low noise level, a normal synchronized mode of A.F.C. operation is established if the initial tuning, Δ_0, does not exceed Δ, i.e., the hold-in range in the given case equals Δ. If the noise level is not very low then, if even the inequality $\Delta_0 \leq \Delta$ holds, a synchronous mode of operation is impossible in the inertialess A.F.C. system under consideration, and the mean frequency of the synchronized generator will differ from the frequency of the fundamental generator.

We now compute the dispersion of the frequency. If we square the right member of equation (9) and average the rapidly varying quantities, we find that

$$(\Delta_0 - \Delta \sin\varphi)^2 + \left(\frac{\Delta}{A_1}\right)^2 [\overline{E_s^2} \sin^2\varphi + \overline{E_t^2} \cos^2\varphi] + \overline{\dot{\psi}^2} =$$
$$= (\Delta_0 - \Delta \sin\varphi)^2 + \sigma^2 + 2\sigma_g^2,$$

since, as a consequence of (4),

$$\overline{E_s^2} \sin^2\varphi + \overline{E_t^2} \cos^2\varphi = \frac{1}{2}\overline{E^2} = \sigma^2, \quad \overline{\dot{\psi}^2} = k_{\dot{\psi}}(0) = 2\sigma_g^2.$$

We may therefore write

$$\overline{\dot{\varphi}^2} = \left(\frac{\Delta}{A_1}\sigma\right)^2 + 2\sigma_g^2 + \int_0^{2\pi} (\Delta_0 - \Delta \sin\varphi)^2 w(\varphi)\, d\varphi. \tag{30}$$

By using formula (25) and computing the integral, we obtain the following result:

$$\overline{\dot{\varphi}^2} = \left(\frac{\Delta}{A_1}\sigma\right)^2 + 2\sigma_g^2 + 2\pi\Delta_0 G -$$
$$- \frac{K_2}{2i}\Delta\, |I_{iD_0}(D)|^{-2} [J_{1-iD_0}(iD)\, J_{iD_0}(iD) - J_{1+iD_0}(iD)\, J_{-iD_0}(iD)],$$

where $J_{1 \pm i\nu}(iz)$ is a Bessel function of a complex order and an imaginary argument.

Consequently, the dispersion of the frequency equals

$$\sigma^2_{\dot{\varphi}} = \overline{\dot{\varphi}^2} - (\overline{\dot{\varphi}})^2 = \left(\Delta \frac{\sigma}{A_1}\right)^2 + 2\sigma^2_{g} + 2\pi G(\Delta_0 - 1) -$$
$$- \frac{K_2\Delta}{2i} | I_{iD_0}(D)|^{-2} [I_{1-iD_0}(D) I_{iD_0}(D) - I_{1+iD_0}(D) I_{-iD_0}(D)], \tag{31}$$

where

$$G = \Delta_0 \frac{\sinh \pi D_0}{2\pi^2 D_0} | I_{iD_0}(D)|^{-2}.$$

SUMMARY

In the analysis of the action of noise on an A.F.C. system, several assumptions were made which permitted the problem to be solved (the phase detector was replaced by a multiplier, the low-frequency filter was assumed to be ideal, the noise was assumed to be a stationary Gaussian function with a spectral density symmetrically distributed about ω_0, etc.). The solution of the problem allowed a number of interesting facts to be established, for example, the presence of phase variations causing divergences of the average frequencies of the fundamental and synchronized generators.

If these assumptions are not made, the mathematical solution of the problem is significantly complicated. Thus, for example, if an actual low-frequency RC-filter is taken into account, the process in the A.F.C. will be described by the following equations:

$$\dot{\varphi} = \Delta_0 - su + \dot{\psi},$$
$$RC \frac{du}{dt} + u = \frac{1}{2} \mu A_2 [(A_1 + E_s) \sin \varphi - E_t \cos \varphi].$$

These equations can be reduced to one nonlinear, second-order differential equation. If condition (11) holds, its solution can then, in theory, be obtained by a special transition from the second-order equation to two first-order equations, with the Fokker-Planck equation then being applied to them [18].

If the noise, n(t), and the fluctuations, ψ(t), are slow, so that the inequality inverse to (11) holds, then a quasi-static approximation may be used [4]. If inequality (11) holds then, for low noise level, independently of its structure, one may use the linearization method in the neighborhood of a stable state.

Received February 23, 1959

LITERATURE CITED

[1] M. R. Kaplanov and V. A. Levin, Automatic Frequency Control [in Russian] Gosénergoizdat (1956).

[2] S. M. Rytov, "Fluctuations and auto-correlations in Tomsonovskii-type systems," [in Russian] J. Exptl.-Theoret. Phys. (USSR), 29, 3 (1955).

[3] V. I. Tikhonov, "The effect of natural fluctuations on self-excited oscillator operation," [in Russian] Vestn. Moscow State Univ. No. 2 (1957).

[4] V. I. Tikhonov and I. N. Amiantov, "The action of slow fluctuations on self-excited oscillators," [in Russian] Radiotekhnika i élektronika 1, 4 (1956).

[5] M. V. Kapranov, "Noise filtering with phase-locked automatic frequency control," [in Russian] Nauchn. dokl. vyssh. shk., Radiotekhnika i élektronika 1 (1958).

[6] L. S. Tyufyakin, "Methods of automatic frequency control," [in Russian] Nauchn. dokl. vyssh. shk. Radiotekhnika i élektronika 1 (1958).

[7] S. G. Margolis, The response of a phase-locked loop to a sinusoid plus noise, Trans. IRE, PGIT, IT-3 (1957).

[8] R. Jaffee and E. Richtin, Design and performance of phase-locked circuits capable of near optimum performance over a wide range of input signal and noise levels. Trans. IRE, PGIT, IT-1 (1955).

[9] V. I. Bunimovich, "Fluctuations in radio receivers," [in Russian] Sovetskoe radio (1951).

[10] P. I. Kuznetsov, R. L. Stratonovich and V. I. Tikhonov, "Correlation functions in the theory of Brownian motion," [in Russian], J. Exptl.-Theoret. Phys. (USSR) 26, 2 (1954).

[11] R. L. Stratonovich, "The synchronization of self-excited oscillators in the presence of noise," [in Russian] Radiotekhnika i élektronika 3, 4 (1958).

[12] S. P. Morgan, Tables of Bessel functions of imaginary order and imaginary argument, Institute of Technology, California, (1947).

[13] S. Ya. Raevskii and R. V. Khokhlov, "On the synchronization of self-excited oscillators by sinusoids in the presence of fluctuating noise," [in Russian], Radiotekhnika i élektronika 3, 4 (1958).

[14] I. G. Akopyan and R. L. Stratonovich, "The establishment of phase synchronization in self-excited oscillators in the presence of fluctuating noise," [in Russian] Nauchn. dokl. vyssh. shk., Ser. fiz.-matem. nauk, No. 1 (1958).

[15] V. I. Tikhonov, "The action of fluctuations on the simplest parametric system," [in Russian], Automation and Remote Control (USSR) 19, 8 (1958).*

[16] L. S. Pontryagin, A. A. Andronov and A. A. Vitt, "On the statistical consideration of dynamic systems," [in Russian] J. Exptl.-Theoret. Phys. (USSR) 3, 3 (1933).

[17] R. V. Khokhlov, "On the theory of locking with external forces of small amplitude," [in Russian] Kokl. AN SSSR 97, 3 (1954).

[18] Ming Chen Wang, and G. E. Uhlenbeck, On the theory of the Brownian motion. Review of modern physics, Vol. 17, No. 2-3 (1945).

* See English translation.

THE EFFECT OF BALANCING LOOP COUPLING ON THE DYNAMIC PROPERTIES OF AC BRIDGES AND COMPENSATORS*

V. Yu. Kneller

(Moscow)

This paper investigates the dependence of the critical gains of the balancing loops of automatic ac bridges and compensators on the amount of coupling of the loops. A knowledge of this dependency is necessary for choosing the tuning of such instruments.

In the design of automatic ac bridges and compensators with two balancing organs (Figs. 1 and 2), these forming an important class of automatic control devices, it is of great practical importance to make the proper choice of the reference voltages' phases for the phase-sensitive indicators (PSI) by means of which the balancing organs form the control signals. As was shown in [1, 2], such devices, in the general case, are two servo systems with mutual coupling, and may be presented by means of the equivalent block schematic shown in Fig. 3. As the PSI reference voltage phases change, the amount of coupling between the balancing loops also changes and this, to a significant degree, changes the system's dynamic properties. In order to provide a tuning of such devices which would guarantee the best dynamic properties, one must know how the magnitude of the coupling between the loops affects the system's dynamics. The present work is devoted to the investigation of this question.

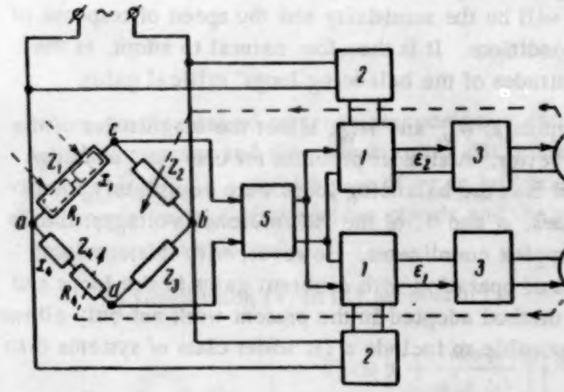


Fig. 1. 1) Phase-sensitive indicator; 2) phase-shifting network; 3) servoamplifier; 4) servomotor; 5) preamplifier.

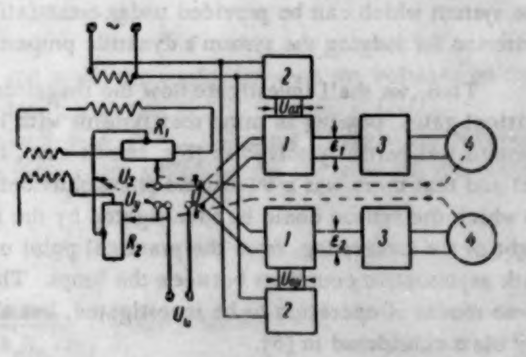


Fig. 2. 1) Phase-sensitive indicator; 2) phase-shifting network; 3) servoamplifier; 4) servomotor.

***This work was presented at the Sixth Scientific-Engineering Conference of the junior members of the IAT AN SSSR on Automatic Control, held January 21, 1959.**

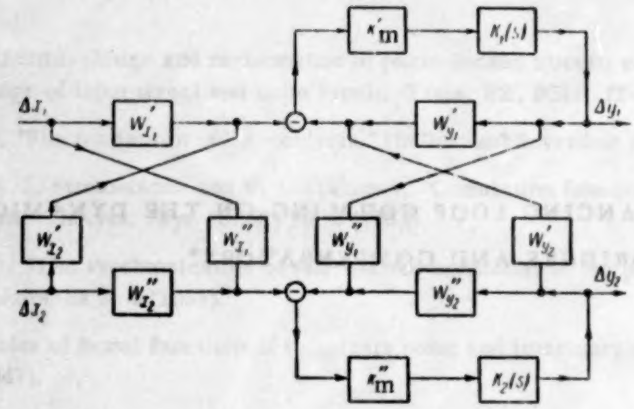


Fig. 3.

We run into similar questions when dealing with the design of systems for controlling objects with coupled quantities and in designing multi-dimensional servo systems [3, 4] but the existing results do not allow an answer to be given for the case we are considering.

Here, we shall consider systems whose block schematic can be represented as in Fig. 3, where the W's are transfer constants whose magnitudes can be chosen at the designer's option. Among such systems are automatic ac bridges with PSI, for any form of bridge scheme, including specially connected balancing organs [5], and also automatic ac compensators [2]. Special cases of such systems are systems with antisymmetric couplings [6].

In choosing a criterion for judging the dynamic properties of the class of systems under consideration, we shall start from the first basic requirements imposed on any automatic measuring device – the requirements of providing a given sensitivity to the quantity to be measured and providing a given speed of response.* The giving of these quantities determines the magnitude of the total gains which must be realized in the balancing loops. With this, the larger the given sensitivity and speed of response, the larger must these gains be. However, stability considerations impose limitations on the size of the balancing loop gains. The critical values of the gains, K_{cr}, are determined by the parameters of the balancing loops and by the magnitudes of the couplings. It is clear that the larger are the magnitudes of the critical gains, the greater will be the sensitivity and the speed of response of the system which can be provided under essentially the same conditions. It is therefore natural to adopt, as the criterion for judging the system's dynamic properties, the magnitudes of the balancing loops' critical gains.

Thus, we shall investigate how the magnitudes of the couplings, W''_{y_1} and W'_{y_2}, affect the magnitudes of the critical gains, bearing in mind that systems with large K_{cr} are better. A similar problem for one class of bridge systems was partially solved in [5]. In this case, it was assumed that the balancing loops were completely identical and that there was a 90-degree phase shift between the phases, φ and θ, of the PSI reference voltages, thanks to which the system could be investigated by the method of complex coordinates. However, with this, one lost sight of the interesting, from the practical point of view, modes of operation with different gains in the loops and with asymmetric couplings between the loops. The analytical method adopted in the present work not only allows these modes of operation to be investigated, but also makes it possible to include a far wider class of systems than the class considered in [5].

With changes in the reference voltage phases, φ and θ, there occur changes both in the coupling coefficients, W''_{y_1} and W'_{y_2}, and in the transfer constants, W'_{x_1} and W''_{x_2} and, consequently, the sensitivity of the instrument to changes in the measured quantities also changes. However, a comparison of systems with different φ and θ for the purpose of finding the system with maximal critical gains must be carried out under essentially identical conditions and, in particular, for one and the same value of sensitivity. Consequently, it is necessary to compare systems

* Here, speed of response is judged, as is customarily done with measuring instruments with automatic balancing, as the time required for a full-scale deflection.

which are obtained from the original one, not only by changing the coupling coefficients, but also by simultaneously changing the gains of parts of the paths from the inputs to the servomotors in such a way that these gains, despite the variations in W'_{x_1} and W''_{x_2} (with changes in φ and θ), remain as they were originally. It is possible to take this into account by replacing our consideration of the system whose block schematic is given in Fig. 3 by consideration of the system whose block schematic is given by Fig. 4. The latter design differs from that of Fig. 3 only in this, that before the motors in the balancing loops there are introduced links with transfer constants $1/k_{var\,1}$ and $1/k_{var\,2}$, where $k_{var\,1}$ and $k_{var\,2}$ are respectively, the factors of W'_{x_1} and W''_{x_2} which change with changes in φ and θ.

Thus, the problem reduces to the investigation of the changes in the critical gains of the loops of the equivalent scheme of Fig. 4 as φ and θ vary. The characteristic equation of this system can be written in the form

$$\left[W'_{y_1}+\frac{k_{var\,1}}{k'_g K_1(s)}\right]\left[W''_{y_2}+\frac{k_{var\,2}}{k''_g K_2(s)}\right]-W'_{y_2}W''_{y_1}=0. \tag{1}$$

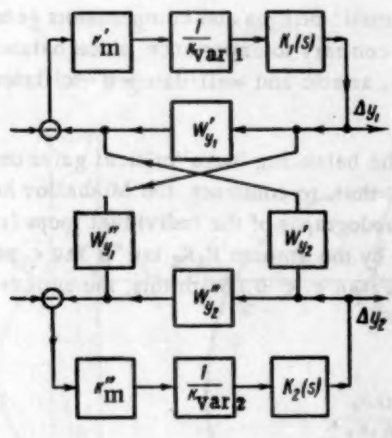


Fig. 4.

In the class of systems under consideration there can be two types of functional dependence of the transfer constants W on φ and θ. In the sequel, we shall analyze these two groups separately.

The first group contains ac bridges with balancing active and passive impedances, placed on one arm, and rectangular-coordinate compensators (Fig. 2).

The second group contains ac bridges with balancing resistors, distributed on different arms (Fig. 1).

Analysis of Systems in the First Group

For systems in the first group, the following relationship is valid [1]:

$$\begin{aligned} W'_{y_1} &= -k\cos\mu, \\ W'_{y_2} &= -k\sin\mu, \quad k_{var} = \cos\mu, \\ W''_{y_1} &= +k\sin\varepsilon, \quad k_{var} = \cos\varepsilon, \\ W''_{y_2} &= -k\cos\varepsilon, \end{aligned} \tag{2}$$

where, in the case of rectangular-coordinate compensators, μ and ε are the angles between the voltages on the balancing organs and the corresponding PSI reference voltages (for the circuit of Fig. 2, $\mu = \angle \dot{U}_x\dot{U}_{0x}$, $\varepsilon = \angle \dot{U}_y\dot{U}_{0y}$). In the case of an ac bridge, $\mu = \varphi + \psi$, $\varepsilon = \theta + \psi - 90°$, where ψ is the angle between the supply voltage of the bridge circuit and the voltage on the bridge arm adjoining the arc to which the controlling organ is connected.

By substituting (2) in (1) we obtain

$$\left[-k\cos\mu+\frac{\cos\mu}{k'_g K(s)}\right]\left[-k\cos\varepsilon+\frac{\cos\varepsilon}{k''_g K_2(s)}\right]+\sin\mu\sin\varepsilon k^2=0.$$

By assuming that $\cos\mu \neq 0$, $\cos\varepsilon \neq 0$, and by using the notation

$$\frac{1}{k'_g k K_1(s)} = -\frac{D_1(s)}{K_1}, \quad \frac{1}{k k''_g K_2(s)} = -\frac{D_2(s)}{K_2},$$

where $K_1 > 0$ and $K_2 > 0$, we obtain

$$[D_1(s)+K_1][D_2(s)+K_2]+K_1K_2\,\mathrm{tg}\,\mu\,\mathrm{tg}\,\varepsilon=0. \tag{4}$$

It is obvious from (4) that a change in the phase of a PSI reference voltage has an effect only on the free term of the characteristic equation. Thus, the problem reduces to the investigation of the effect of changes in the characteristic equation's free term on the system's stability. For given $D_1(s)$, $D_2(s)$, K_1 and K_2 this is easily carried out by using the well-known methods of automatic control theory. However, our problem is to obtain a series of general recommendations which will be valid for the class of systems under consideration for different values of these parameters. These recommendations must give the possibility of choosing φ and θ from the point of view of providing a maximal K_{cr} without having recourse each time to concrete numerical calculations. For the problem thus posed we impose the following limitations which take into account the properties of the class of systems under consideration:

1) we shall assume that both balancing loops have approximately identical hodographs, i.e., $D_1(j\omega) \approx D_2(j\omega) = D(j\omega)$;

2) we shall assume that the Mikhailov hodographs of the balancing loops take the form of unwinding spirals which do not intersect themselves.

The first limitation is valid insofar as both balancing loops of automatic bridges and compensators generally are made up of mono-typic elements. The second limitation is also not contrary to experience, since balancing loops of such automatic measuring devices ordinarily consist of aperiodic, astatic and well-damped oscillatory links, and with this the hodograph $D(j\omega)$ does not intersect itself [7].

We now elucidate the qualitative character of the dependence of the balancing loops' critical gains on μ and ϵ, in considering them as Mikhailov hodographs. It is clear from (4) that, to construct the Mikhailov hodograph of the complete system, it is necessary to multiply the Mikhailov hodographs of the individual loops (curve 1, Fig. 5,a) and thereafter to translate the resulting hodograph (Fig. 5,b) by the amount $K_1K_2 \tan \mu \tan \epsilon$ parallel to the real axis, to the right if $\tan \mu \tan \epsilon > 0$ and to the left if $\tan \mu \tan \epsilon < 0$. With this, the system remains stable if (Fig. 5, a and b),* for $\tan \mu \tan \epsilon > 0$,

$$|\operatorname{tg}\mu \operatorname{tg}\varepsilon| < \frac{OA_1}{K_1K_2} \text{ and } |\operatorname{tg}\mu \operatorname{tg}\varepsilon| < \frac{OA_3}{K_1K_2}, \tag{5}$$

and, for $\tan \mu \tan \epsilon \leq 0$,

$$|\operatorname{tg}\mu \operatorname{tg}\varepsilon| < \frac{OA_2}{K_1K_2} \text{ and } |\operatorname{tg}\mu \operatorname{tg}\varepsilon| < 1. \tag{6}$$

Here we shall consider only those portions of the Mikhailov hodographs of the balancing loops which include the points a_1, a_2 and a_3 since, by virtue of the second assumption made, it is precisely these points which form the points A_1, A_2 and A_3 which are situated the closest to point 0. As is obvious from (5) and (6), the values of the balancing loops' gains are bounded by the quantity $\tan \mu$ and ϵ. If the lengths of the segments OA_i ($i = 1, 2, 3$) did not depend on K_1K_2, then a single answer would derive from (5) and (6), namely, the larger $\tan \mu \tan \epsilon$, the smaller the critical gains of the balancing loops. But the lengths of the OA_i also depend on K_1K_2 and it is therefore necessary to explain to which side the points A_i are translated when K_1K_2 increases.

We assume initially that $K_1 = K_2 = K$. With this, as K increases, segment OA_2 always decreases, and, consequently, OA_2 decreases, i.e., point A_2 is translated to the left.

Thus, the right member of inequality (6) decreases as K increases and we obtain an unequivocal answer, namely, that for $\tan \mu \tan \epsilon < 0$, as μ and ϵ increase the critical gains decrease, and this means that if it is desired to increase K it is then necessary to choose $\tan \mu \tan \epsilon$ as small as possible. The law of change of OA_1 with increasing K depends on the form of the Mikhailov hodograph. Thus, for example, if the hodograph has the form of curve 2, Fig. 5,a then, with certain limits of K's values, as K increases the segment OA_1 can increase. However, for further increases of K, beginning with a certain value of it, OA_2 again decreases with increasing K.

* On Figs. 5-12, the left side of the figure (labeled "a") represents the Mikhailov hodographs of the individual loops, while the right side (labeled "b") represents the product of the Mikhailov hodographs of the individual loops.

Thus, we arrive at the conclusion that also for $\tan\mu\tan\epsilon > 0$, if it is desired to increase K then $\tan\mu\tan\epsilon$ should be chosen as small as possible.

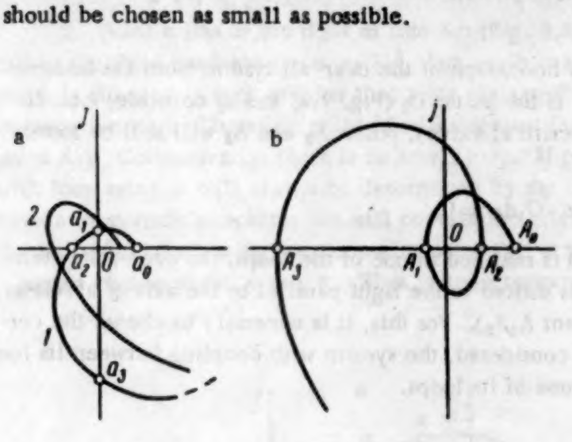


Fig. 5.

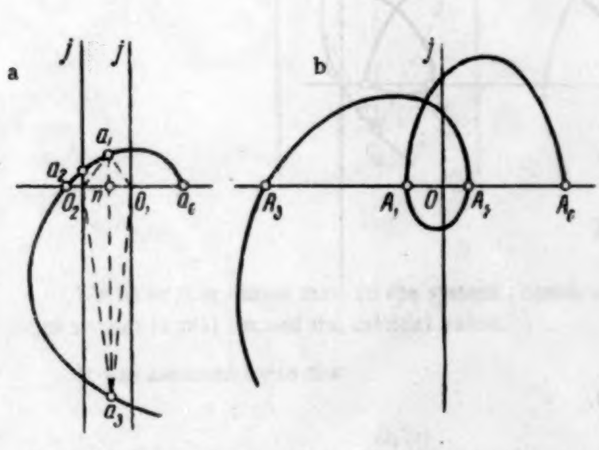


Fig. 6.

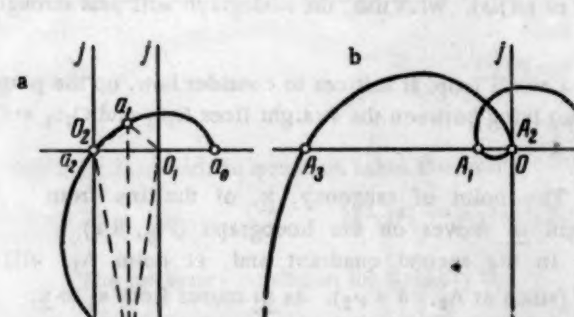


Fig. 7.

It is not difficult to see that, for K close to K_{cr}, when K increases there is a decrease both in OA_2 and also in OA_1 and OA_3, while for $K = K_{cr}$, as OA_2 reduces to zero so will at least one of the quantities OA_1 or OA_3, i.e., the points A_i go to zero from two sides simultaneously. Consequently, for $K = K_{cr}$, the over-all system cannot be made stable for any values of μ and ϵ.

Thus, for the case considered we obtain the unequivocal answer: as the coupling of the balancing loops increases, the loops' critical gains decrease. The maximum possible K_{cr} in the loops are realized when the system is completely uncoupled.

We now consider the case when $K_1 \neq K_2$. We show the Mikhailov hodographs of both balancing loops on one figure (Fig. 6,a). The balancing loops with different gains are represented on Fig. 6,a in such wise that the axes of the hodographs' ordinates do not coincide. As the gain of one of the loops increases, the corresponding axis of ordinates on Fig. 6,a is translated to the left. Here, $OA = O_1a_1 \cdot O_2a_2$, and with an increase in gain of one of the loops, for example, K_1, segment OA_2 is decreased. Consequently, for $\tan\mu\tan\epsilon < 0$, the larger is $|\tan\mu\tan\epsilon|$, the smaller is the critical value of K_1K_2.

The frequencies for which the hodograph

$$M(j\omega) = [D_1(j\omega) + K_1][D_2(j\omega) + K_2]$$

crosses the real axis at the points A_1 and A_3 (Fig. 6,b) are found on Fig. 6,a as the frequencies at the points of intersection of the individual loop hodographs with the straight line parallel to the axes of ordinates going through the point n which bisects the segment O_2O_1.

We now consider how OA_1 and OA_3 change as the gains in the balancing loops change. Here, several cases are possible. If the gains simultaneously increase in such fashion that their difference remains invariant, then OA_1 and OA_3, as in the previously considered case, decrease starting with some values of K_1 and K_2. Therefore, also in the given case, for increasing K_{cr}, $|\tan\mu\tan\epsilon|$ should be decreased. However, if only one of the gains increases, the law of change of OA_1 and OA_3 is determined by two phenomena: the decreases in the quantities na_1 and na_3 as one of the axes of ordinates is translated to the left, and the change in the magnitude of O_1O_2.

Thus, if the lesser of the loop gains is increased, there occurs a simultaneous decrease both of a_1n and a_3n and of O_1O_2. Consequently, $OA_1 = O_1a_1 \cdot O_2a_1$ and $OA_3 = O_1a_3 \cdot O_2a_3$ decrease as rapidly as in the previous case.

Therefore it is again desirable to decrease $|\tan \mu \tan \epsilon|$. If now the larger of the gains is increased (for example, K_2 on Fig. 6,a) then, simultaneously with a decrease in na_1 and na_3, there occurs an increase of O_1O_2 and a consequence of this can be an increase of OA_1 and OA_3.

In considering the process of constructing the Mikhailov hodograph of the over-all system from the hodographs of the individual systems, we may note the following. If the points O_2 (Fig. 7,a) and a_2 coincide, i.e., in the case when $OA_2 = O$ (the gain in one of the loops equals a critical value), points A_2 and A_3 will still be found to the right of point O, since

$$O_2a_2 \cdot a_1O_1 \neq 0 \text{ and } O_2a_3 \cdot O_1a_3 \neq 0.$$

This means that, although the critical value of the gain is realized in one of the loops, the over-all system can be made stable if the hodograph $M_1(j\omega)M_2(j\omega)$ (Fig. 7,b) is shifted to the right parallel to the axis of abscissas in such fashion that point O falls inside the loop (inside segment A_1A_2). For this, it is necessary to choose the corresponding magnitude of $\tan \mu \tan \epsilon > 0$. Thus, in the case considered, the system with coupling between its loops permits a gain greater than the critical one to be realized in one of its loops.

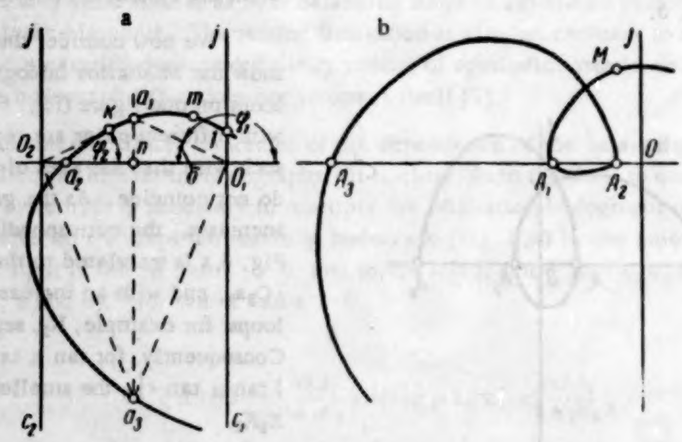


Fig. 8.

We now explain to what value it is possible to increase the gain in one of the balancing loops while simultaneously so choosing the value of $\tan \mu \tan \epsilon$ that the system remains stable. The system can be made stable, at the cost of changing $\tan \mu \tan \epsilon$, if the hodograph $M(j\omega) = M_1(j\omega)M_2(j\omega)$ has a loop inside of which it is possible to place point O at the cost of only one parallel transfer of $M(j\omega)$. With this, the hodograph will pass through the necessary number of quadrants.

To determine the conditions for which $M(j\omega)$ will have a small loop, it suffices to consider how, on the plane of hodograph $M(j\omega)$, the points of hodographs $M_1(j\omega)$ and $M_2(j\omega)$ lying between the straight lines O_1c_1 and O_2c_2 are imaged. We consider the possible variations.

1. Point a_2 lies to the left of point $\underline{n}$ (Fig. 8,a). The point of tangency, $\underline{k}$, of the line from O_2 to the hodograph lies to the left of a_1n. As the point $\underline{m}$ moves on the hodograph (Fig. 8,a) from point 1 to point a_1, point M (Fig. 8,b) will move in the second quadrant and, at point A_1, will fall on the real axis, since $\varphi = \varphi_1 + \varphi_2 = 180° - \delta + \varphi_2 = 180°$ (since at A_2, $\delta = \varphi_2$). As $\underline{m}$ moves from a_1 to $\underline{k}$, point M moves into the third quadrant, since $\varphi = 180° - \delta + \varphi_2$ but $\varphi_2 > \delta$ and φ_2 increases with increasing ω. After point $\underline{k}$, φ_2 begins to decrease and, consequently, point M approaches the real axis, falling on it at point A_2. $OA_2 < OA_1$, since $O_1a_1 \cdot O_2a_1 > O_1n \cdot O_2n$. As $\underline{m}$ moves from a_2 to a_3, $|\varphi_2| > |\delta|$ but, since $\varphi = 180° + |\delta| - |\varphi_2|$, M is translated into the second quadrant, then falls upon point A_3 on the real axis. $OA_3 > OA_2$, since $O_1a_3 \cdot O_2a_3 >$ $> O_1a_2 \cdot O_2a_2$. With subsequent motion of $\underline{m}$ from point a_3, $|\delta| > |\varphi_2|$, and the variation of φ is determined by the variation of $|\delta|$. The latter quantity increases with increasing ω.

Thus, for $O_2a_1 > O_2n$ and $O_2k_1 > O2n$, there will always be a small loop which intersects the real axis in such fashion that, by translation of $M(j\omega)$, one can always place point O inside the loop.

2. Point k lies to the right of line a_1n (Fig. 9,a). As m moves from a_1 to a_2, point M (Fig. 9,b) does not fall in the third quadrant since $\varphi_2 > \delta$, but φ_2 decreases as m moves from a_1 to a_2. As ω increases, the hodograph is changed in such manner that after passage through point a_2 angle φ_2 increases in modulus, but retains its negative sign. Therefore point M still does not fall in the third quadrant after its passage through point A_2 (to point A_3). Consequently, there is no small loop. If point n is found to the left of point a_1 then the motion of M with increasing ω will always be determined by the character of the variations of δ and the hodograph $M(j\omega)$ will have a monotonic character but will contain the origin of coordinates, O, only so many times as the hodograph of the individual balancing loop contains point O_1. It is also impossible to make such a system stable by means of a proper choice of $\tan\mu \tan\epsilon$. Thus, a loop appears only in the case when point k lies to the left of line a_1n.

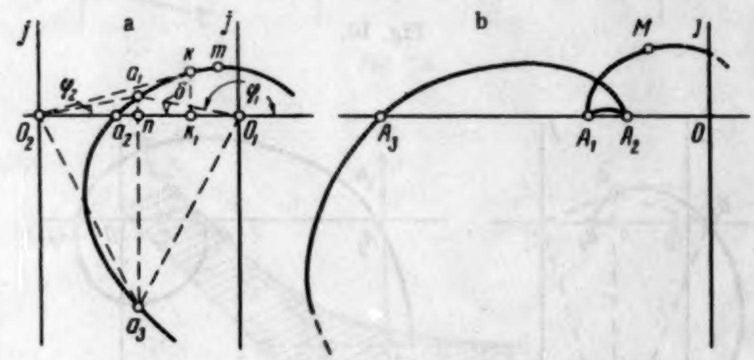


Fig. 9.

We have thus shown that, in the system considered, it is possible to choose the gain for one of the balancing loops so that it will exceed the critical value.

It was assumed by us that

$$\frac{D_1(s)}{K_1} = -\frac{1}{k_g'kK_1(s)}, \quad \frac{D_2(s)}{K_2} = \frac{1}{k_g'kK_2(s)}.$$

We now change the polarity in one of the balancing loops, for example, in the first loop. Then,

$$\frac{D_1(s)}{K_1} = \frac{1}{k_g'kK_1(s)} \quad (K_1 > 0)$$

and the characteristic equation takes the form:

$$[D_1(s) - K_1][D_2(s) + K_2] - K_1K_2 \operatorname{tg}\mu \operatorname{tg}\varepsilon = 0. \tag{7}$$

The necessary condition for stability is

$$\tan\mu \tan\epsilon < -1.$$

As before, we shall assume that $D_1(j\omega) = D_2(j\omega) = D(j\omega)$, and we shall consider several cases.

1. $K_1 > K_{cr}$, $K_2 > K_{cr}$, $O_1n < O_1k_1$.

By proceeding as we did above, constructing hodograph $M(j\omega)$ from the hodographs of the individual systems and using the analogous reasoning, we can show that, in this case, there will be a segment of the axis of abscissas which will be contained by the hodograph $M(j\omega)$ the necessary number of times for stability (on Fig. 10,b, this segment is A_0A_1).

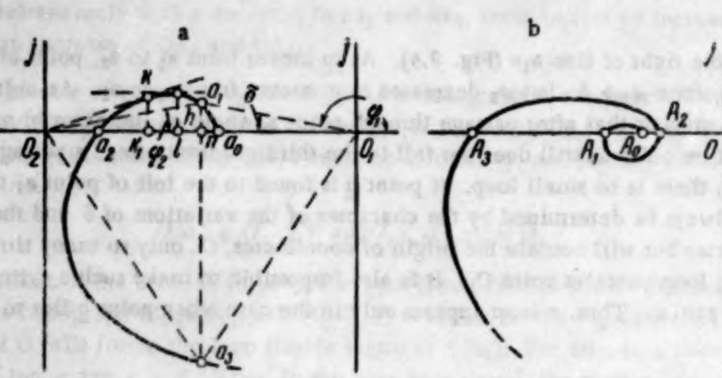


Fig. 10.

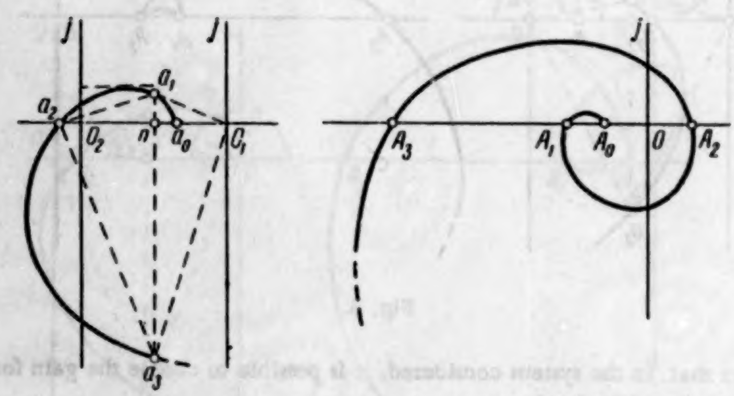


Fig. 11.

In order that the system be stable, it is necessary so to choose the magnitude of tan μ tan ϵ that point O will lie on the interior of this segment.

$$OA_0 < K_1 K_2 |\operatorname{tg}\mu \operatorname{tg}\varepsilon| < OA_1 \tag{8}$$

or

$$K_1 K_2 < K_1 K_2 |\operatorname{tg}\mu \operatorname{tg}\varepsilon| < h^2 + \left(\frac{K_1 + K_2}{2}\right)^2. \tag{9}$$

Thus, in the case under consideration, it is possible, by choosing the proper magnitude of $|\tan \epsilon \tan \mu| \neq 0$, to make the system stable for arbitrarily large gains in the individual loops, if these gains satisfy the conditions

$$lO_1 > \frac{K_1 + K_2}{2} > K_1, \tag{10}$$

where l is the projection on the axis of abscissas of the hodograph point with the maximum positive ordinate (located on the first turn).

2. $K_2 < K_{cr}$, $O_1 n > K_1$.

As follows from what has been constructed (Fig. 11, a and b), in this case the system can also be made stable by choosing the proper value of tan μ tan ϵ

$$A_2 O > |\operatorname{tg}\mu \operatorname{tg}\varepsilon| K_1 K_2 > K_1 K_2.$$

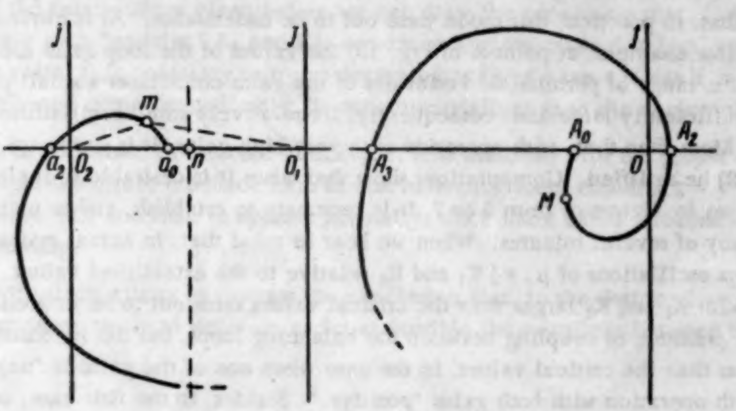


Fig. 12.

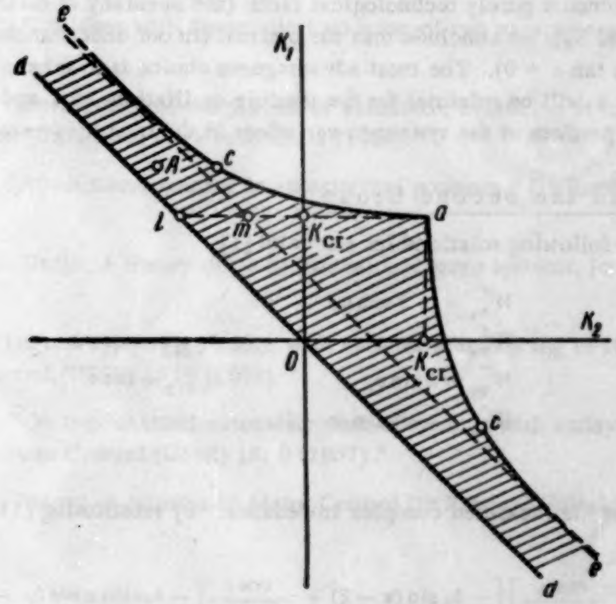


Fig. 13.

3. $K_2 < K_{cr}$, $O_1n < K_1$.

As point m moves from point O ($\omega = 0$) (Fig. 12,a) the phase angle immediately begins to increase and point M (Fig. 12,b) moves into the third quadrant. With this, a small loop is not formed, and the system cannot be made stable by properly choosing $\tan \mu \tan \epsilon$.

The results obtained can be presented graphically on one figure (Fig. 13) if the signs before K_1 or K_2 in (4) or (7) are taken to refer to K_1 or K_2 themselves and, in correspondence with this, if the values of K_1 and K_2 are laid off on the negative or positive semiaxes in the K_1K_2 plane. The hatched region in Fig. 13, excluding the boundaries and the coordinate axes, is the region of values of K_1 and K_2 for which the system can be made stable at the cost of properly choosing $\tan \mu \tan \epsilon$. Boundary c-a-c of this region, for large K, asymptotically approaches the line e-e which is parallel to line d-d which bounds the region on the left. The magnitude of lm depends on the form of $D(j\omega)$ and is proportional to a_0l (Fig. 10,a). It is clear from Fig. 13 that the system can be made stable for arbitrarily large gains in the balancing loops if these gains are related by (10). For that, it is necessary to choose the magnitude of $\tan \mu \tan \epsilon$ which satisfied (9). Thus, the maximum gains are realized in coupled systems.

However, system operation with gains exceeding the critical values is attended by a number of disadvantages which are so essential that, in practice, this mode turns out to be undesirable. As is obvious from Fig. 13, with operation in this mode (for example, at point A of Fig. 13) the values of the loop gains are bounded both above and below. With this, the range of permissible variations of the gains constitutes a small part of their absolute values, when they are sufficiently large and, consequently, for relatively small oscillations of the gains the system can become unstable. More than that, with operation with very high gains, it is necessary that μ and ϵ be strictly constant in order that (9) be satisfied. Computations show that when it is desirable to realize loop gains which exceed the critical values by factors of from 5 to 7, it is necessary to establish, and to maintain, the values of μ and ϵ with an accuracy of several minutes. When we bear in mind that, in actual systems of the class considered, there are always oscillations of μ, ϵ, K_1 and K_2 relative to the established values, we are led to the conclusion that operation with K_1 and K_2 larger than the critical values turns out to be impossible in practice. Thus, the drawback is not the presence of coupling between the balancing loops, but the inconstancy of the couplings. Operation with gains less than the critical values, in the case when one of the gains is "negative," has no superiority in comparison with operation with both gains "positive." Besides, in the first case, one of the gains is also bounded below. Therefore, it is advantageous to operate in the first quadrant of Fig. 13 with "positive" gains less than the critical value. With this, the maximal loop gains (close to K_{cr}) are realized in coupled systems.

Thus, by taking into account purely technological facts (the accuracy of establishing μ and ϵ and the possible oscillations of K_1 and K_2), we conclude that the optimal (in our understanding) system turns out to be the uncoupled system ($\tan\mu\tan\epsilon = 0$). The most advantageous choice is to take $\mu = 0$ and $\epsilon = 0$ since, with this, the value of $\tan\mu\ \tan\epsilon$ will be minimal for the possible oscillations in μ and ϵ. Moreover, with this the sensitivity of the measuring portions of the system to variations in the measured parameters will be maximal.

Analysis of Systems in the Second Group

For these systems, the following relationships are valid [1]:

$$\begin{aligned} &W'_{\psi_1} = -k\cos\mu, & & \\ &W'_{\psi_2} = k_3\cos(\mu-\xi), & &k_{\text{var}1} = \cos\mu, \\ &W''_{\psi_1} = k\sin\varepsilon, & &k_{\text{var}2} = \cos\varepsilon, \\ &W''_{\psi_2} = -k_3\sin(\varepsilon-\xi), & & \end{aligned} \tag{11}$$

where ξ is the phase angle of the measured complex impedance. By substituting (11) in (4) we obtain

$$\left[-k\cos\mu + \frac{\cos\mu}{k'_g K_1(s)}\right]\left[-k_3\sin(\varepsilon-\xi) + \frac{\cos\varepsilon}{k''_g K_2(s)}\right] - k_3 k\sin\varepsilon\cos(\mu-\varepsilon) = 0. \tag{12}$$

By assuming that $\cos\mu \neq 0$, $\cos\epsilon \neq 0$, and using the notation

$$\frac{1}{kk'_g K_1(s)} = -\frac{D_1(s)}{K_1}, \quad \frac{1}{k_3 k''_g K_2(s)} = \frac{D_2(s)}{K_2},$$

where $K_1 > 0$, $K_2 > 0$, we obtain

$$[K_1 + D_1(s)]\left[K_2\frac{\sin(\xi-\varepsilon)}{\cos\varepsilon} + D_2(s)\right] + K_1K_2\frac{\cos(\mu-\xi)}{\cos\mu}\operatorname{tg}\varepsilon = 0. \tag{13}$$

From a comparison of (4) and (13) it is clear that, if we consider the term $K_2\frac{\sin(\xi-\varepsilon)}{\cos\varepsilon}$ as the gain of the second balancing loop, denoting it by K_2', then all the previously made conclusions relative to the magnitudes of the gains in the balancing loops which can be realized with the proper choices of coupling coefficients remain in force for bridges of the second type as well. All remarks previously made anent the choice of the quantity $K_1K_2\tan\mu\tan\epsilon$ now relate to the magnitude of $K_1K_2\frac{\cos(\mu-\xi)}{\cos\mu}\tan\epsilon$.

On the basis of the relationships given before we can draw the conclusion that, in the given case also, it is advantageous to operate with "positive" K_1 and K_2', not exceeding the critical values, and that to realize, with this, the greatest loop gains, it is necessary to try to decrease $[\cos(\mu - \xi)\tan\epsilon]/\cos\mu$ as far as possible, i.e., to uncouple the system. Whence come immediately the recommendations as to the choice of ϵ and μ.

It is interesting to note that, in the case considered, it is possible, with the proper choice of the coupling coefficients, to uncouple the direct feedback loop in one balancing loop, choosing $\xi = \epsilon$ [cf. (13) and Fig. 3] and, with this, the bridge will continue to operate normally, since there exists a second coupling via the first loop and the inter-loop coupling.

Thus, the analysis given allows us to draw the conclusion that, in the design of automatic ac bridges and compensators, it is necessary to try to decrease as far as possible the couplings between the balancing loops.

Received March 30, 1959

LITERATURE CITED

[1] V. Yu. Kneller, "AC bridges with automatic balancing of two parameters, part II," [in Russian] Otchet IAT AN SSSR (1957).

[2] V. Yu. Kneller, "Convergence of the process of automatic balancing of measuring circuits with two balancing organs," [in Russian] Izmeritel'naya tekhnika 3 (1959).*

[3] Yu. G. Kornilov, "Autonomous control as an extremal problem," [in Russian] Izv. AN SSSR, OTN 4 (1954).

[4] M. Golomb and E. Usdin, A theory of multidimensional servo systems, Journal of Franklin Inst., vol. 253, No. 1, (1952).

[5] V. Yu. Kneller, "On one type of ac bridge with automatic balancing of two parameters," [in Russian] Automation and Remote Control (USSR) 19, 2 (1958).*

[6] A. A. Krasovskii, "On two-channel automatic control systems with antisymmetric couplings," [in Russian] Automation and Remote Control (USSR) 18, 2 (1957).*

[7] M. A. Aizerman, Theory of Automatic Motor Control [in Russian] Gostekhizdat (1952).

* See English translation.

EXPERIMENTAL DETERMINATION OF AUTOMATIC CONTROL SYSTEM LINKS' TRANSFER FUNCTIONS BY MEANS OF STANDARD ELECTRONIC MODELS

L. N. Darovskikh

A method is suggested for determining the coefficients of transfer functions by means of standard electronic models (analogs). A table is provided which allows one to set up a system of linear algebraic equations for computing the coefficients of the transfer function with, or without, a polynomial in the numerator, and also when there is only a polynomial in the numerator. Examples of the application of this method are given.

The existing methods for determining the transfer function of a link, or of a set of links, of an automatic control system (ACS) [1-7] require a great amount of processing of experimental results, or are applicable only in special cases [1-3]. The method presented here is characterized by a minimum number of computations subsequent to the experimentation, reducing to the solution of a system of linear algebraic equations.

Posing of the Question

The idea of our method consists in the transformation of the transfer function

$$W(p) = \frac{B(p)}{A(p)}, \tag{1}$$

where

$$A(p) = 1 + a_1 p + a_2 p^2 + \ldots + a_n p^n, \quad B(p) = 1 + b_1 p + b_2 p^2 + \ldots + b_m p^m,$$

by means of a constructed scheme which includes a solving block of standard electronic analogs. This transformation renders it possible to obtain a system of linear algebraic equations for the computation of the unknown coefficients of transfer function (1).

The transformation has the goal of obtaining a new transfer function with constant terms in the numerator, which will be a linear function of the coefficients, a_i and b_i, of transfer function (1). The size of a constant term is determined by the indications of a voltmeter connected to the output as the gain of the circuit obtaining the investigated portion of the ACS.

Setting $m < n$, we consider the expression

$$W_1(p) = -\frac{1}{p}[W(p) - 1]. \tag{2}$$

We now carry out the transformation:

$$W_1(p) = -\frac{(b_1 - a_1) + (b_2 - a_2)\,p + (b_3 - a_3)\,p^2 + \ldots + (-a_n)\,p^{n-1}}{A(p)}.$$

A circuit with such a transfer function has the gain

$$k_1 = -(b_1 - a_1).$$

A circuit with the transfer function

$$W_2(p) = -\frac{1}{p}\,[W_1(p) - k_1]$$

has the gain

$$k_2 = k_1 a_1 - a_2 + b_2, \text{ etc.}$$

The construction of a circuit for carrying out the transformation cited presents no difficulties.

If the transfer function contains a gain k which differs from unity then, with the introduction of a nomial unit input signal, $1^* = 1/k$, the link's transfer function can be considered as a transfer function with unit gain. Therefore, the gain k does not enter into consideration; the nominal unit, 1^*, will be called simply the unit input signal.

Determination of the Time Constant of the Simplest Transfer Function

Let a link of an ACS be characterized by the simplest transfer function

$$W(p) = \frac{1}{1 + Tp}. \tag{3}$$

In this case, expression (2) gives

$$W_1(p) = \frac{T}{1 + Tp}.$$

When a unit step input voltage is applied to a circuit with transfer function $W_1(p)$, a voltage equal to T is established at the output. The operations of algebraic addition of $W(p) - 1$ and multiplication by $-1/p$ can be implemented by means of operational amplifiers in the set of standard analogs.

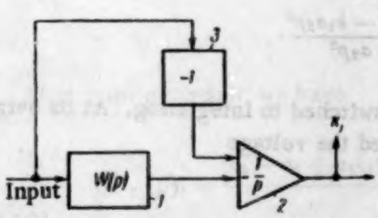


Fig. 1.

Figure 1 shows the block schematic for determining the time constant, T, of a link with transfer function (3). To the circuit's input there is applied a voltage of 1 volt (in the general case, m volts, where m is the scale factor).

The output of the link under investigation is connected to the input of operation amplifier 2 which operates in the summing mode. By means of compensator 3, a null voltage is established at the output of amplifier 2. Verification of the compensation of the signal at the output of the investigated link is implemented by switching block 2 to integration.

When zero is attained at the output of block 2, the input signal is taken off. Block 2 is switched to integrating. A unit step voltage is applied to the circuit's input. Obviously, a voltage is established at the integrating block's output equal, in volts, to

$$k_1 = T. \tag{4}$$

Determination of the Coefficients of Transfer Functions of the Form $W(p) = 1/(1 + a_1p + a_2p^2)$

Figure 2 shows the block schematic for finding the coefficients of the transfer function

$$W(p) = \frac{1}{1 + a_1p + a_2p^2}. \tag{5}$$

For this there are required two operational amplifier blocks and two compensators.

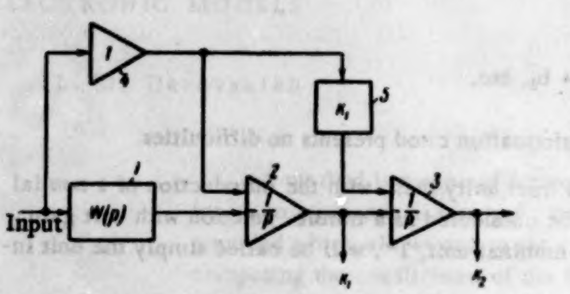


Fig. 2.

A one-volt signal is applied to the circuit's input. A voltage is established at its output, also equal to 1 volt. Block 2 is in the summing mode. At its output, by means of compensator 4, a null voltage is established, which corresponds to the passage of the input signal through a link with transfer function

$$W_1(p) = \frac{1}{1 + a_1p + a_2p^2} - 1 = -\frac{p(a_1 + a_2p)}{1 + a_1p + a_2p^2}.$$

The input signal is then taken off and block 2 is switched to integrating. When a unit step voltage is applied to the circuit's input, a voltage is established at the output of integrating block 2 which corresponds to the passage of the input signal through a link with transfer function

$$W_2(p) = -\frac{p(a_1 + a_2p)}{1 + a_1p + a_2p^2}\left(-\frac{1}{p}\right) = \frac{a_1 + a_2p}{1 + a_1p + a_2p^2}$$

and which equals

$$k_1 = a_1. \tag{6}$$

This voltage is applied to the input of block 3, which is in the summing mode. By means of compensator 5, a null voltage is established at the output. This corresponds to the passage of the input signal through a link with transfer function

$$W_3(p) = \frac{a_1 + a_2p}{1 + a_1p + a_2p^2} - k_1 = \frac{(a_2 - a_1k_1)\,p - k_1a_2p^2}{1 + a_1p + a_2p^2}.$$

Then the voltage is at the circuit's input is taken off and block 3 is switched to integrating. At its output, when a unit step voltage is applied to the circuit's input, there is established the voltage

$$k_2 = k_1a_1 - a_2, \tag{7}$$

corresponding to the passage of the input signal through a link with transfer function

$$W_4(p) = -\frac{(a_2 - a_1k_1) - k_1a_2p}{1 + a_1p + a_2p^2}.$$

Equations (6) and (7) form the system

$$k_1 = a_1,\ k_2 = k_1a_1 - a_2,$$

whose solution gives the values of the unknown coefficients of the transfer function:

$$a_1 = k_1,\ a_2 = k_1^2 - k_2.$$

Consequently, for determining the coefficients a_1 and a_2 of a given transfer function (5), only two voltage measurements need be made. The subsequent work reduces to the solution of linear algebraic equations.

Obtaining the System of Algebraic Equations for Determining the Transfer Function Coefficients in the General Case

Let

$$W(p) = \frac{1 + b_1 p + b_2 p^2 + \ldots + b_m p^m}{1 + a_1 p + a_2 p^2 + \ldots + a_n p^n} = \frac{B(p)}{A(p)}. \tag{8}$$

In solving the problem of obtaining a system of algebraic equations for determining the coefficients of a transfer function of the type cited, we shall understand by addition, and by multiplication by $-1/p$, the operations which were cited out in the foregoing samples.

Letting $m < n$, we consider the expression

$$W_1(p) = -\frac{1}{p}[W(p) - 1].$$

We carry out the transformation:

$$W_1(p) = -\frac{(b_1 - a_1) + (b_2 - a_2)p + (b_3 - a_3)p^2 + \ldots}{A(p)}$$
$$\frac{\ldots + (b_m - a_m)p^{m-1} + \ldots + (-a_n)p^{n-1}}{A(p)}.$$

The constant term, $(b_1 - a_1)$, in the numerator is determined as the circuit's gain, k_1, when the transfer function is $w_1(p)$:

$$-k_1 = b_1 - a_1. \tag{9}$$

We now consider the expression

$$W_2(p) = -\frac{1}{p}[W_1(p) - k_1].$$

After transformation, we have

$$W_2(p) = \frac{(b_2 - a_2 + a_1k_1) + (b_3 - a_3 + a_2k_1)p + \ldots + (b_m - a_m + a_{m-1}k_1)p^{m-2}}{A(p)} + \ldots$$
$$\ldots + \frac{(-a_n + a_{n-1}k_1)p^{n-2} + a_n k_1 p^{n-1}}{A(p)}.$$

The constant term in the numerator is determined as the gain, k_2, of the circuit with transfer function $W_2(p)$:

$$k_2 = b_2 - a_2 + a_1 k_1. \tag{10}$$

We next consider the expression

$$W_3(p) = -\frac{1}{p}[W_2(p) - k_2].$$

We carry out the transformation:

$$W_3(p) = -\frac{(b_3 - a_3 + a_2k_1 - a_1k_2) + (b_4 - a_4 + a_3k_1 - a_2k_2)\,p + \ldots}{A(p)}$$

$$\frac{\ldots + (b_m - a_m + a_{m-1}k_1 - a_{m-2}k_2)\,p^{m-3}}{A(p)} - \ldots$$

$$\ldots - \frac{(-a_n + a_{n-1}k_1 - a_{n-2}k_2)p^{n-3} + (a_nk_1 - a_{n-1}k_2)\,p^{n-2} - a_nk_2\,p^{n-1}}{A(p)}.$$

The constant term in the numerator is determined as the gain, k_3, of the circuit with transfer function $W_3(p)$:

$$-k_3 = b_3 - a_3 + a_2k_1 - a_1k_2. \tag{11}$$

By noting the regularity of the formation of the coefficients k_i, and by applying mathematical induction to equations (9-11), we can set up the following system of algebraic equations:

for m < n

$$\begin{gathered}
k_1 + b_1 = a_1,\\
k_2 - b_2 = k_1a_1 - a_2,\\
k_3 + b_3 = k_2a_1 - k_1a_2 + a_3,\\
k_4 - b_4 = k_3a_1 - k_2a_2 + k_1a_3 - a_4\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_m + (-1)^{m-1}b_m = k_{m-1}a_1 - k_{m-2}a_2 + k_{m-3}a_3 - \ldots\\
\ldots + (-1)^m k_1a_{m-1} + (-1)^{m-1}a_m,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_n = k_{n-1}a_1 - k_{n-2}a_2 + k_{n-3}a_3 - \ldots + (-1)^n k_1a_{n-1} + (-1)^{n-1}a_n,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_{n+m} = k_{n+m-1}a_1 - k_{n+m-2}a_2 + \ldots + (-1)^{n+m}k_{m+1}a_{n-1} + (-1)^{n+m-1}k_ma_n
\end{gathered} \tag{12}$$

for m > n

$$\begin{gathered}
k_1 + b_1 = a_1,\\
k_2 - b_2 = k_1a_1 - a_2,\\
k_3 + b_3 = k_2a_1 - k_1a_2 + a_3,\\
k_4 - b_4 = k_3a_1 - k_2a_2 + k_1a_3 - a_4,\\
k_5 + b_5 = k_4a_1 - k_3a_2 + k_2a_3 - k_1a_4 + a_5.\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_n + (-1)^{n-1}b_n = k_{n-1}a_1 - k_{n-2}a_2 + k_{n-3}a_3 - \ldots\\
+ (-1)^n k_1a_{n-1} + (-1)^{n-1}a_n,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_m + (-1)^{m-1}b_m = k_{m-1}a_1 - k_{m-2}a_2 + k_{m-3}a_3 - \ldots\\
+ (-1)^m k_{m-n+1}a_{n-1} + (-1)^{m-1}k_{m-n}a_n,\\
k_{m+1} = k_ma_1 - k_{m-1}a_2 + k_{m-2}a_3 - \ldots\\
+ (-1)^{m-1}k_{m-n+2}a_{n-1} + (-1)^m k_{m-n+1}a_n,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_{m+n} = k_{m+n-1}a_1 - k_{m+n-2}a_2 + \ldots\\
+ (-1)^{m+n}k_{m+1}a_{n-1} + (-1)^{m+n-1}k_ma_n.
\end{gathered} \tag{13}$$

By taking as shown the coefficients $k_1, k_2, k_3, \ldots$ and by solving system (12), for $m < n$, or system (13), for $m > n$, each of these systems consisting of $n + m$ algebraic equations, we can find the unknown coefficients of the transfer function of the ACS link under investigation. The coefficients $k_1, k_2, k_3, \ldots$ are determined experimentally by the methodology explained above.

If it was necessary to have two compensators and two operational amplifiers for the determination of the transfer function coefficients when there was only a second-degree polynomial in the denominator, then in the case of $m + n$ unknown coefficients of transfer function (8) it is required to have $m + n$ compensators and $m + n$ operational amplifiers. The circuit in which this apparatus is connected remains essentially as it was before.

We now consider two particular cases of formula (8).

Case 1. $B(p) = 1$. The system of algebraic equations has the form:

$$\begin{aligned} k_1 &= a_1, \\ k_2 &= k_1 a_1 - a_2, \\ &\cdots\cdots\cdots \\ k_n &= k_{n-1} a_1 - k_{n-2} a_2 + \ldots + (-1)^n k_1 a_{n-1} + (-1)^{n-1} a_n. \end{aligned} \tag{14}$$

In considering this system of equations, we can draw the conclusion that, generally speaking, in the determination of the coefficients of a transfer function with no polynomial in the numerator, a knowledge of the order of the polynomial in the denominator is not necessary, since the coefficient a_1 is obtained by means of the first measurement of voltage, the coefficient a_2 is obtained after the second voltage measurement, a_3 is obtained after the third measurement, etc. Thus, the process of experimentation can also elicit the order of the denominator's polynomial (of course, within the limits of experimental accuracy). The method for determining the coefficients when there is a polynomial in the transfer function's numerator necessitates either previous knowledge or eliciting of the degrees of the polynomials in the numerator and denominator, which must be considered a disadvantage of the method presented here.

Case 2. $A(p) = 1$. The system of algebraic equations has the form:

$$\begin{aligned} k_1 + b_1 &= 0, \\ k_2 - b_2 &= 0, \\ k_3 + b_3 &= 0, \\ &\cdots\cdots\cdots \\ k_m + (-1)^{m-1} b_m &= 0. \end{aligned}$$

The case when the transfer function has no polynomial in the denominator is of great practical interest, since by measuring the voltages $k_1, k_2, k_3, \ldots$ it is possible to determine in the most direct way all the unknown coefficients of the transfer function. From this the result derives that it is advantageous in experimenting to replace a link with transfer function $W_1(p) = 1/A(p)$ by a link with transfer function $W_2(p) = A(p)$ which can be implemented by fully automatic means.*

For convenience in using the method presented here for determining the transfer function, we provide a table below which allows one to set up the system of algebraic equations for the determination of the coefficients of the transfer function of the link under consideration, independently of whether or not the operator p is contained in the numerator or denominator of the transfer function. The table is made up for transfer functions containing $m + n = 10$ unknown coefficients. The basis of this was the determinant set up from the system of algebraic equations in (13). We now consider two examples of the use of the table.

Example 1. It is required to set up the system of algebraic equations for the transfer function of the form

$$W(p) = \frac{1}{1 + a_1 p + a_2 p^2 + a_3 p^3}.$$

* It is well known that an operational amplifier of an electronic analog which is shunted by a negative feedback path which is a link with transfer function $K(p)$ has the transfer function $W(p) \approx 1/K(p)$.

The transfer function contains three unknowns; consequently, it is necessary to set up a system of three equations. For setting up this system, only three rows of the table are required. We obtain one equation from each row, resulting in the system

$$a_1 = k_1,$$
$$k_1a_1 - a_2 = k_2,$$
$$k_2a_1 - k_1a_2 + a_3 = k_3.$$

Example 2. The transfer function has the form:

$$W(p) = \frac{1 + b_1p + b_2p^2}{1 + a_1p + a_2p^2 + a_3p^3}.$$

The number of unknowns is five. Consequently, for setting up the system it is necessary to use five rows of the table:

from the first row

$$a_1 = k_1 + b_1. \tag{16}$$

from the second row

$$k_1a_1 - a_2 = k_2 - b_2. \tag{17}$$

from the third row

$$k_2a_1 - k_1a_2 + a_3 = k_3. \tag{18}$$

from the fourth row

$$k_3a_1 - k_2a_2 + k_1a_3 = k_4. \tag{19}$$

from the fifth row

$$k_4a_1 - k_3a_2 + k_2a_3 = k_5. \tag{20}$$

Equations (16-20) form a system of algebraic equations whose solution gives the unknown coefficients of given transfer function (15).

For setting up the systems of algebraic equations, m + n rows of the table were used. The equations set up from the following rows reduce mathematically to identities. For the purpose of checking the experimental results it is advantageous to introduce still another equation, the (m + n + 1)'st. For transfer function (15), the check equation is

$$k_5a_1 - k_4a_2 + k_3a_3 = k_6.$$

To obtain coefficient k_6 it is necessary to connect one additional operational amplifier in the experimental setup. In practice, the control equation must hold within the limits of accuracy of the given experiment. If the left member of the equation differs too sharply from the right member, it is necessary to repeat the experiment more carefully, using other input voltages for this.

Errors in Practical Application of the Method

The accuracy of the determination of the unknown coefficients of transfer functions depends on the magnitude of the zero drift of the operational amplifiers, on the error of the integration process and also on the error of the measuring instruments (the voltmeters).

The error due to operational amplifier zero drift depends on the type of dc integrator and, in many types of electronic analogs, is practically zero. The error is small and can easily be taken into account. The error in the parameters found by the method presented will depend, to the greatest degree, on the accuracy of the voltage

measurements by the dc voltmeters. This accuracy will, in practice, not be greater than 0.5%. With an increase in the number of unknown coefficients, the number of working units also increases, and some error is unavoidable.

Since the errors will have a fundamentally random character, repetition of the experiment several times and averaging the obtained values of the coefficients can increase the accuracy of the results, and their reliability can be estimated by statistical processing. Verification of the degree of identity of the check equation might serve as a sufficient criterion for judging the accuracy of the results.

Received June 30, 1958

LITERATURE CITED

[1] A. A. Voronov, Elements of Automatic Control Theory [in Russian] Voenizdat (1954).

[2] M. A. Aizerman, Theory of Automatic Motor Control [in Russian] Gostekhizdat (1952).

[3] R. Ol'denburg and G. Sartorius, Automatic Control Dynamics, [in Russian] Gosénergoizdat (1949).

[4] Ya. Z. Tsypkin, "Determination of the dynamic parameters of systems described by differential equations of not higher than second order, by means of oscillograms of the transient responses," [in Russian], in the collection "Trudy Vsesoyuznogo zaochnogo énergeticheskogo instituta." Élektrotekhnika 6 (1955).

[5] G. L. Rabkin, B. A. Mitrofanov and Yu. O. Shterenberg, "On determining the numerical values of the coefficients of transfer functions of linearized links in systems with experimental frequency characteristics," [in Russian] Automation and Remote Control (USSR) 16, 5 (1955).

[6] M. P. Simoyu, "The determination of the coefficients of the transfer functions of linearized links in auto-control systems," [in Russian] Automation and Remote Control (USSR) 18, 6 (1957).*

[7] A. A. Kardashov and L. V. Karnyushin, "Determining system parameters from experimental (given) frequency characteristics," [in Russian] Automation and Remote Control (USSR 19, 4 (1958).*

* See English translation.

MINIMIZATION OF THE BOOLEAN FUNCTIONS CHARACTERIZING SWITCHING CIRCUITS

M. A. Gavrilov

(Moscow)

The paper considers a method for minimizing the Boolean functions which characterize switching circuits, the method being based on an analysis of the conditions to be realized by the operation of the switching circuits.

One of the essential problems of the modern structural theory of relay (switching) devices is the problem of relay circuits minimization, i.e., decreasing the number of elements in them. Since switching circuits can be characterized by Boolean functions, the aforementioned problem coincides with the problem of minimizing the latter but, however, has its own specific peculiarities.

A large number of works have been devoted to questions of minimization. These questions were already touched upon to a significant extent by Claude Shannon in 1938 [1] with the publication of one of the first works in the domain of switching theory. The first regular method was described in [2] in the form of the so-called "minimizing charts." Work in the minimization domain progressed considerably with the publication of the papers of W. Quine [3, 4] who first formulated the minimization problem in its present form. A number of subsequently published works [5-9] were based on Quine's ideas. The majority of them consider the problem of minimizing Boolean functions as the problem of finding those of their terms which are "neighbors" [10], i.e., which contain some letter both nonnegated and negated. By using the well-known relationship

$$fx + f\bar{x} = f(x + \bar{x}) = f, \tag{1}$$

and also by successively applying the well-known operations

$$x + \varphi(x, y, \ldots, w) = x + f(0, \varphi, \ldots w) \tag{2}$$

and

$$f_1x + f_2\bar{x} + f_1f_2 = f_1x + f_2\bar{x}, \tag{3}$$

to the terms of the Boolean functions, one reduces the number of terms in the functions and the number of letters in the terms until the point when the operations just cited can no longer be applied. The Boolean functions thus obtained are considered minimal.

In the works of various authors [10-13], various methods have been suggested for placing the terms of Boolean functions in such fashion that it becomes possible to determine graphically the terms' "neighbors." There have been described various types of tables and their transformations, based on relationship (1), subsidiary means in the form of minimizing charts or stencils allowing one to find and compute the number of "neighbors," etc. An original method of minimizing was suggested by K. K. Voishvillo [14] which solves the minimization problem by finding the so-called "closing sets." Most of the works cited deal with the problem of minimizing the structure

Fig. 1.

of nonsequential switching devices and do not deal with the minimization of sequential switching devices although, as will be shown in the sequel, taking the special features of sequential devices into account significantly simplifies the minimization procedure.

In the present paper we consider a somewhat different approach to the solution of the minimization problem, namely, an approach based on the analysis of the conditions of operation and nonoperation which the switching circuit is to realize. The same approach is applied both to sequential and nonsequential circuits, and makes it possible to significantly simplify the process of minimization itself.* We understand by the latter, as do a number of other authors, the problem of obtaining, based on the conditions of operation of a device as described in the form of a switching table or by Boolean functions given in normal form, the same form but now containing the least possible number of letters.

A special feature of Boolean functions which characterize switching circuits is that, in contradistinction to ordinary Boolean functions, they may have added to them certain nominal terms, deriving from the operations specified for the switching device. To explain this, we consider in more detail the operation of a switching circuit.

In its generalized form the function of a switching device is to transform a combination of discrete stimuli applied to its input to a definite combination of these same stimuli at its output (Fig. 1).

For nonsequential switching circuits, i.e., those in which no facility is provided for any temporal sequence of stimuli application at the input or obtaining a stimuli at the output, the circuit's conditions of operation can be completely given if each combination of stimuli applied to the input is put into correspondence with a definite combination of stimuli at the output.

We shall denote the presence of a stimulus (for example, the closing of a circuit, or the appearance of a high voltage) by unity, and the absence of a stimulus (for example, the opening of a circuit, or the presence of a low voltage) by zero. Then, the conditions of operation of a switching circuit can be given in the form of a table of binary numbers (the so-called "switching table") in which, to each binary number which corresponds to a combination of stimuli at the inputs or, as we shall say in the sequel, a combination of input "states," there is set up on the same row a binary number which corresponds to the combination of stimuli at the outputs (the combination of output "states").

Such a table of binary numbers is shown in Fig. 2,a for a particular case of the conditions of operation of a switching circuit with four inputs and two outputs. For each of the circuit's outputs all of the combinations of the input states may be divided into two groups: one which corresponds to ones at the output and the other corresponding to zeroes at the output. Figures 2b and 2c give the tables of binary numbers for each of the outputs, Z_1 and Z_2, separately.

We shall call the combination of the input states which correspond to ones at a given output as its "working" states, and shall denote the set of them by the symbol F_1. The working states define the conditions of output operation (conditions of circuit closure, or the appearance of a high voltage). They can be described by a Boolean function whose terms characterize the working states at the inputs. Such a Boolean function is easily obtained if, for example, we write each of the combinations of working states (inverted for a zero, not inverted for a one) in the form of a product of letters corresponding to the circuit's inputs and then take the sum of such products. Thus, for example, for the operating conditions of the table in Fig. 2 we shall have:**

---

* What is presented below is a generalization, development, and the mechanical basis of, work previously published by the author on the minimization question. The method described, and its application to sequential circuits, was first mentioned in [15]. It was then further developed, as applied to the same devices, in [16] and, finally, as applied to nonsequential circuits, was mentioned in the author's paper, "Developmental work on the structural theory of switching circuits in the Soviet Union," given at Harvard University in April, 1957 at the International Symposium on Switching Theory. In a generalized form, this method was presented by the author at a joint seminar on switching theory of the IAT AN SSSR, the Laboratory for scientific problems of line communications, and the Mathematical Inst. AN SSSR, in April, 1958.

** Below, as is done in many works on the structural theory of switching circuits, the conjunction sign is replaced by the sign for algebraic multiplication, and the disjunction sign by the sign for algebraic addition.

for element Z_1

$$F_{1(z_1)} = \bar{a}\bar{b}\bar{c}d + \bar{a}\bar{b}cd + ab\bar{c}d + abcd; \tag{4}$$

for element Z_2

$$F_{1(z_2)} = \bar{a}\bar{b}cd + ab\bar{c}\bar{d} + ab\bar{c}d. \tag{5}$$

It is easily seen that, with such an algorithm for making the transition from the switching table to an analytic form of writing the operating conditions of a relay device, one obtains expressions in the form of Boolean functions written in the so-called "complete normal disjunctive form."

A	B	C	D	Z_1	Z_2
0	0	0	0	0	0
0	0	0	1	1	0
0	0	1	1	1	1
1	0	0	1	0	0
1	1	0	0	0	1
1	1	0	1	1	1
1	1	1	1	1	0

a

	A	B	C	D	Z_1
1	0	0	0	1	1
2	0	0	1	1	1
3	1	1	0	1	1
4	1	1	1	1	1
	0	0	0	0	0
	1	0	0	1	0
	1	1	0	0	0

b

	A	B	C	D	Z_2
1	0	0	1	1	1
2	1	1	0	0	1
3	1	1	0	1	1
	0	0	0	0	0
	0	0	0	1	0
	1	0	0	1	0
	1	1	1	1	0

c

Fig. 2.

The second group of conditions, namely, those combinations of states which correspond to values of zero for the given output of the switching circuit, will be denoted in their entirety by the symbol F_0, and will be called "forbidden." Just as for the working states, the Boolean function characterizing the forbidden states will be obtained by writing each of the forbidden states in the form of a product of letters, corresponding to the circuit's inputs, and then taking the sum of these products.

Since, as was shown in [15, 16], the operating conditions of the output can be expressed in terms of the forbidden states by taking the inverse of the Boolean function characterizing the latter, the Boolean function for the combinations of the forbidden states, if used in writing analytically the conditions of operation, will be obtained in the so-called "complete normal conjunctive form."

For a complete grasp of the operating conditions of a switching circuit, it is necessary to know the nature of the remaining possible combinations of input states which, when added to F_1 and F_0, make up the total of 2^n combinations (where $\underline{n}$ is the total number of inputs). These combinations can be divided into two categories. One of these is that of the combinations known to be impossible for them to exist, i.e., for no conditions of operation of the objects supplying stimuli to the switching circuit can they occur. Obviously, any term of the Boolean function corresponding to one of these states will, when written in the disjunctive form, always have the value zero, and when written in the conjunctive form, the value one. We shall call such combinations of states "unused," and will denote the set of them by F_n.

To the other category of states belong those of which it is known that it is indifferent for the operating conditions of the switching device whether a one does or does not appear at the output when they appear at the input. Such combinations of states will be called "indifferent," and the set of them will be denoted by F_p.

It is obvious that terms of a Boolean function which characterize any unused or indifferent combinations of states can be added to the Boolean function characterizing the set of working combinations of states without disturbing the operating conditions of the switching circuit: when the Boolean function is written in disjunctive form, these terms will enter in as factors.

The addition of terms corresponding to unused combinations of conditions will not change the initial switching table, since each such term equals zero, and the addition of terms corresponding to indifferent combinations of states will give one of the permissible variants of the switching table. The term of the Boolean function which correspond to unused or to indifferent states are the nominal terms which were mentioned earlier.

All further reasoning will apply to Boolean functions written in the normal disjunctive form called, for brevity, simply the "normal form." However, all the minimization methods considered will apply equally when the operating conditions are expressed in terms of Boolean functions written in the normal conjunctive form, it being necessary only to interchange F_1 with F_0 and to interchange zeroes and ones.

When Boolean functions are written in the normal disjunctive form, all sets of combinations of input states corresponding to the operating conditions of some output of the switching device can be written in the form:

$$F_{(Z)} = F_i + [F_n + F_p], \tag{6}$$

where the terms in the brackets are nominal ones, i.e., those which can, at will, be included or not in the Boolean function characterizing the operating conditions.

It is clear that the switching table is realizable if, for each of the outputs, no combination of input states is simultaneously in F_1 and in F_0.*

We shall say that a Boolean function, written in complete normal form, "realizes" some one of the combination of input states of a switching circuit, corresponding to the conditions of operation, or nonoperation, of some circuit output, if the combination of letters in one of the terms of this Boolean function, when the noninverted (i.e., nonnegated) letters are replaced by ones and the inverted letters are replaced by zeroes, is contained in the row of the switching table which corresponds to this combination.

For a Boolean function written in normal form, a combination of input states will be "realized" by one of its terms if, when the aforementioned replacement of letters by zeroes and ones is made, the combination is contained in part of the row, or the complete row, of the switching table which corresponds to this combination. We note that a term of a normal Boolean function can realize simultaneously several combinations of input states.

By the "length," l, of a term in the normal form of a Boolean function we shall mean the number of letters entering into it, and by the "length," L, of the normal form itself we shall mean the total number of letters in it. Then, the process of minimizing a Boolean function or, as we shall say in the sequel, of obtaining its "minimal form" can be reduced to decreasing the length of its normal form. It is obvious that such a decreasing of length can be carried out only up to the point where the Boolean function still corresponds to the operating conditions of the switching circuit, i.e., will not realize any of the combinations of states entering into F_0.

Thus, the minimal form of a Boolean function is that normal form of it which realizes all the combinations of states in F_1 and does not realize any combination of states in F_0, whereby a decrease in length of any of its terms by at least one letter, excluding its complete expunging, would lead to a condition wherein the Boolean function would realize at least one combination of states which is simultaneously in F_1 and F_0.

It is obvious that if a Boolean function is written in complete normal form its terms f will have maximum length. Let these terms lie in some set $M\{2^n\}$. Then, the operating conditions of some output Z, including unused and indifferent states, can be presented as some subset of the set M, $N \subset M$. With this

$$F_{(Z)} = N = \sum f_i + \left[\sum f_n + \sum f_p\right], \tag{7}$$

where

$$\sum f_i = F_i, \quad \sum f_n = F_n \text{ and } \sum f_p = F_p.$$

*This condition for realizability was formulated somewhat differently in [16].

It is obvious that any term φ of the normal form of a Boolean function will have a length l_φ equal to or less that the length, l_f, of the terms of the complete normal form: $l_\varphi \leq l_f$.

We shall say that a Boolean function in either its normal or its complete normal form, $F_{(Z)} = \sum \varphi$ or $F_{(Z)} = \sum f$, realizes the operating conditions of output Z if the totality of its terms realizes all the working states and realizes none of the forbidden states.

We shall consider that a term of the normal form of a Boolean function has minimal length, or is "(minimal" (φ_{min}), if it contains such a number of letters that decreasing it by at least one letter (excluding the very last one) would lead to a condition where this term realizes at least one combination of states which is simultaneously in $F_{1(Z)}$ and in $F_{0}(Z)$.

Theorem 1. The minimal form of a Boolean function which realizes the operating conditions of one of the outputs Z of a switching circuit is that form which contains only minimal terms.*

$$F_{(Z)\min} = \sum \varphi_{\min}. \tag{8}$$

Indeed, if at least one of the terms φ has a length l greater than the minimal length, then this means that the length L of the expression $F_{(Z)}$ can be decreased. If at least one of the terms φ has a length less than the minimal length, then this term will simultaneously realize at least one of the states in $F_{1(Z)}$ and $F_{0(Z)}$, i.e., the normal form of the Boolean function containing it will not, by our definition, by minimal.

On the basis of Theorem 1, we can present the following algorithm for obtaining the minimal form of a Boolean function realizing the condition $F_{1(Z)}$: we shall choose terms of the normal form which realize states of $F_{1(Z)}$, starting with terms containing only one letter and, gradually increasing the number of letters, we shall verify, for each of these terms, whether or not they realize some state from $F_{0(Z)}$. The shortest terms φ which do not realize states in $F_{0(Z)}$ will be entered in the minimal form of the Boolean function which realizes $F_{1(Z)}$."**

We note that each of these minimal forms will automatically include the necessary number of terms of the complete normal function from f_1, f_n and f_p.

We now consider the application of this algorithm to the example given above in Fig. 2, starting this consideration with the operating conditions for output Z_1. If we consider for this the elements of f_0, we immediately convince ourselves that c = 1 is absent from them. Therefore, the minimal term $\underline{c}$ can enter into the minimal form of the Boolean function which realizes $F_{1(Z_1)}$. It is obvious that all terms containing this letter as a factor will not be minimal, and these, therefore, need not be tried in the remainder of the process. The two-letter combinations which do not give realizations of states in F_0 are the terms $\bar{a}d$ and bd.

Thus, the minimal form of the Boolean function which realizes the operating conditions of output Z_1 will have the form:

$$F_{(Z_1)} = c + \bar{a}d + bd. \tag{9}$$

By determining the minimal form of the Boolean function which realizes the operating conditions for output Z_2 in the same manner, we obtain

$$F_{(Z_2)} = \bar{a}c + a\bar{d} + \bar{b}c + b\bar{c} + b\bar{d}. \tag{10}$$

---

* A theorem analogous in its formulation was proven by W. Quine in [3]. However, in their essentials, the theorem stated above and Quine's theorem are different, and from them there flow different algorithms for obtaining minimal forms of Boolean functions. Quine's theorem is based on the concept of the so-called "prime implicants" which are obtained with the use of formulae (1) and (3). A special procedure is thereby necessitated for taking the unused states into account (Cf., [4]). Theorem 1, formulated above, is based on the concept of "minimal terms" which are obtained by the method, previously given by the author, of analyzing the conditions of realizability of a switching table (Cf., [16]), which automatically takes into account the necessary unused states.

** As applied to nonsequential circuits, this algorithm was first used by the author in the previously cited paper given at Harvard University. It was subsequently applied in [17] for the case of minimizing the number of diodes in structures with diode nets.

A minimal form which contains all the minimal terms will be called a "general" minimal form. Such general minimal forms are those given by expressions (9) and (10). In many cases, however, the conditions f_i can be realized by means of some, not necessarily all, the minimal terms. Obviously, it is sufficient for this that there be enough such terms to realize all the conditions in f_i without exception. We shall call such a minimal form of a Boolean function which realizes all the operating conditions of some output of a switching circuit, and which contains the minimal necessary number of minimal terms, a "particular" minimal form.

To determine, from a general minimal form, all the possible particular forms, it is convenient to use the methodology presented by McCluskey [6].* In this methodology, a table is set up for the general minimal form, wherein it is noted, for each of the minimal terms, which of the individual states of F_i, from the switching table, it realizes relative to the given output of the switching circuit.

f_i	$Q = c$	$P = \bar{a}d$	$R = bd$
1		+	
2	+	+	
3			+
4	+		+

a

f_i	$T = \bar{a}c$	$P = a\bar{d}$	$Q = \bar{b}c$	$R = b\bar{c}$	$S = b\bar{d}$
1	+		+		
2		+		+	+
3				+	

b

Fig. 3.

We number the individual conditions for $F_{i(Z_1)}$ and $F_{i(Z_2)}$ as indicated on Fig. 2, we denote the minimal terms by capital letters and we denote by a cross the fact that an individual condition has been realized, this cross being placed at the intersection of the row representing this condition and the column representing the corresponding minimal term. Then, for the example considered above with the two outputs Z_1 and Z_2, we obtain the tables shown in Fig. 3. It is obvious that, from the entire set of minimal terms contained in the general minimal form, it is necessary, for each of the partial minimal forms, to choose those of their combinations for which the conditions of F_i would be realized without exception.

We write this in the form of formulae, wherein we use the logical connective "or" to denote the conditions for realizing the individual f_i and the logical connective "and" to denote the condition, cited above, of the necessity of realizing all the f_i.

Then, by using the conventional notation for conjunction and disjunction, we obtain:

for the table of Fig. 3,a:

$$\Phi_1 = P\,(Q \vee P)\,R\,(Q \vee R); \tag{11}$$

for the table of Fig. 3,b

$$\Phi_2 = (T \vee Q)\,(P \vee R \vee S)\,R. \tag{11a}$$

By using the well-known formula for simplifying Boolean functions

$$xf(x, y, \ldots, w) = xf(1, y, \ldots, w) \tag{12}$$

and by transforming the formulae obtained to the normal disjunctive form, we shall have

*This problem can also be solved by means of the method of closing sets, suggested by K. Volshvillo.

$$\Phi_1 = P(Q \vee P)R(Q \vee R) = PR, \tag{13}$$

$$\Phi_2 = (T \vee Q)(P \vee R \vee S)R = (T \vee Q)R = TR \vee QR. \tag{14}$$

In correspondence with this, the minimal forms of the Boolean functions for the outputs Z_1 and Z_2 are written in the following way:

$$F_{(Z_1)\min} = ad + bd, \tag{15}$$

$$F_{(Z_2)\min} = \bar{a}c + b\bar{c} = \bar{b}c + b\bar{c}. \tag{16}$$

The method just presented gives a general solution to the problem of minimizing the Boolean function characterizing a switching circuit, but requires the execution of a sorting operation which, in certain cases, can be quite awkward.

If it is necessary to obtain only one of the partial minimal forms, then the sorting operation is not necessary. To determine the algorithm for obtaining such forms, we consider Theorem 2.

Theorem 2. A Boolean function, expressed in the normal form F' which has one or several terms φ' of length l' less than the minimal length, i.e., which realize simultaneously, not only all the f_1, but also certain f_0, can be taken to a normal form which realizes only f_1 if, for the conditions f_0 mentioned above, one sets up one of the partial minimal forms ψ, and then takes the product F' $\overline{\Psi}$.

A term of minimal length is formed from a certain set of terms of f_1, f_n and f_p of the complete normal form as the result of employing operation (1). If its length is less than the minimal length, this means that at least one f_0 enters into the set of terms f forming it.

$$\varphi' = \sum\nolimits_\alpha f_1 + \sum\nolimits_\beta f_n + \sum\nolimits_\gamma f_p + \sum\nolimits_\delta f_0. \tag{17}$$

To remove this contradiction, it suffices to multiply φ' by such a function that the resulting product goes to zero when there is present any of the combinations of input states corresponding to each of the f_0 entering into $\sum_\delta f_0$. In the most general case, this can be the inverse of $\sum_\delta f_0$.

However, in correspondence with the definitions given above, this multiplier can also be inverse of any normal form ψ of the Boolean function which satisfies the condition: ψ realizes $\sum_\delta f_0$.

Thus,

$$\varphi = \varphi'\overline{\psi} = \left(\sum\nolimits_\alpha f_1 + \sum\nolimits_\beta f_n + \sum\nolimits_\gamma f_p + \sum\nolimits_\delta f_0\right)\overline{\psi}. \tag{18}$$

These arguments are also valid for the entire function F', since F' = $\sum \varphi'$.

Thus, indeed,

$$F_1 = F'\overline{\Psi}. \tag{19}$$

The theorem is thereby proven.

We note that by carrying out the operations corresponding to Theorem 2 we can obtain, in addition to the terms realizing f_1, terms which only realize f_n and f_p. One can convince oneself of this by multiplying out the parentheses in (18):

$$\varphi = \overline{\psi}\sum\nolimits_\alpha f_1 + \overline{\psi}\left(\sum\nolimits_\beta f_n + \sum\nolimits_\gamma f_p\right). \tag{20}$$

The second term of this expression can give terms which do not enter into the set of minimal terms for f_1, but which only realize some terms of f_n and f_p. Therefore, the solution obtained in this case must be compared with the conditions for f_1, and those terms which do not realize the latter must be discarded.

On the basis of Theorem 2 we can advance an algorithm which allows one to obtain particular minimal forms without the use of sorting.

It is necessary for this to take a minimal set of certain terms φ' with lengths less than the minimal length (for example, containing only one letter) and which realize the conditions of f_1 without duplication, such that each of them realizes a minimal number of conditions of f_0, and then carry out on these terms the operations cited above.

Let

$$F' = \varphi_1' + \varphi_2' + \ldots + \varphi_n'. \tag{21}$$

For each φ' we determine the corresponding function ψ and then carry out the termwise multiplication of each φ' by the inverse, $\overline{\psi}$:

$$F_1 = \varphi_1' \overline{\psi}_1 + \varphi_2' \overline{\psi}_2 + \ldots + \varphi_n' \overline{\psi}_n. \tag{22}$$

The function obtained must be the minimal form of the Boolean function which realizes all the conditions of f_1.

In the general case, each of the ψ is a sum of minimal terms ξ realizing F_0. If any of these terms has a length greater than the minimal length, i.e., has more letters composing it, or if one of the expressions ψ has a superfluous term, i.e., ψ is not a particular minimal form, then the corresponding product, $\varphi_k'\overline{\psi}_k$ will form a superfluous term of φ or a term of φ with a superfluous letter, i.e., which is not minimal. Thus, in both these cases, the Boolean function obtained will either not be a minimal form or not be a particular minimal form.

Now, let a term ξ with a length less than minimal enter into one of the expressions ψ_k. The function ψ_k will then realize, not only terms of f_0, but also some terms of f_1. This means that the corresponding product, $\varphi_k'\psi_k$, will go to zero for some combinations of input states, corresponding to f_1 and, since the φ' were so chosen by us that they realize all the conditions f_1 without duplication, this then means that F_1 will not realize all conditions f_1, i.e., will not be a minimal form of the Boolean function.

We now illustrate the application of the algorithm considered here by the example of the switching table of Fig. 2.

It can easily be seen for output Z_1 (Fig. 2,b) that all the f_1 can be realized by means of the term $\varphi' = d$, whereby this term also realizes a minimal number of f_0. We now set up the table of conditions of f_1 and f_0 realized by $\varphi' = d$ (Fig. 4,a). It is easily seen that the minimal term which satisfies the condition that it only realizes F_0 will be $\xi = a\overline{b}$. Thus, $\psi = a\overline{b}$ and

$$F_1 = \varphi'\overline{\psi} = d\,\overline{(a\bar{b})} = \bar{a}d + bd,$$

which corresponds with the result obtained previously.

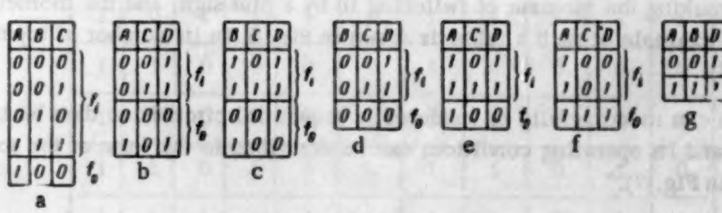


Fig. 4.

We now take, as the φ', the terms a and $\overline{b}$, each of which realizes two of the f_0. The corresponding tables for conditions f_1 and f_0 are shown in Fig. 4,b and c. As may easily be noted, for $\overline{b}$ the conditions of f_0 are realized by the terms ξ, equal to $\overline{d}$ and a, and for a these conditions are realized by the ξ equal to $\overline{b}$ and $\overline{d}$. Therefore,

$$\varphi_{\bar{b}} = \bar{b}\,\overline{(\bar{d} + a)} = \bar{a}\bar{b}d \text{ and } \varphi_a = a\,\overline{(\bar{b} + \bar{d})} = abd.$$

The terms obtained are not minimal, since $\overline{b}$ and $\underline{a}$ do not satisfy the conditions formulated above.

If we now choose, as the terms φ', the terms $\overline{a}$ and $\underline{b}$, each of which realizes only one f_0 (Fig. 4, d and e), we obtain

$$\varphi_{\bar{a}} = \bar{a}\bar{\bar{d}} = \bar{a}d \text{ and } \varphi_b = b\bar{\bar{d}} = bd.$$

In this case, since the terms φ' were chosen in correspondence with the conditions formulated for the Boolean function realizing f_1, they are minimal.

For the output Z_2, the conditions formulated for the φ' are satisfied by terms $\underline{b}$ and $\underline{c}$. By analyzing the corresponding tables of conditions f_1 and f_0 (Fig. 4, f and g), we obtain

$$\varphi_b = b(\bar{c}) = b\bar{c} \text{ and } \varphi_c = c(\bar{b}) = c\bar{b}.$$

The algorithm being considered is not universal, since the results obtained with its use depend heavily on the proper choice of the terms φ'. However, even when their choice is not completely satisfactory, the algorithm significantly speeds up the process of determining the particular minimal forms, since the solution obtained is close to the minimal solution and can always be quickly taken to the minimal solution by testing each of the terms obtained by means of the first algorithm presented here.

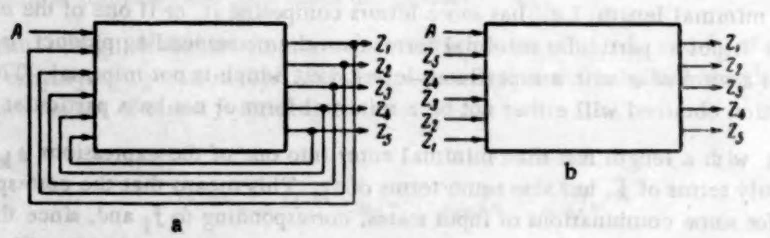


Fig. 5.

We note that the algorithms which derive from Theorems 1 and 2 can be applied, not only when the conditions F_1' and F_0 are given in the form of a switching table, but also even when they are expressed by Boolean functions in normal form. On the basis of these latter formulae, one can set up tables of binary numbers in which the letters lacking in the corresponding terms of the normal forms are noted by means of lines. After this the determination of the minimal terms is carried out by means of the algorithms described above, but with it being taken into account that a line can mean either a zero or a one.

We now turn to the consideration of sequential switching circuits. These circuits differ from the nonsequential ones in that they contain connections between the output states and the input states. This property is a consequence of the fact that for part (or for all) of the output there is a feedback path to the inputs (Fig, 5,a).

It is convenient, in this case, to present the switching table in the form of a sequence of combinations of circuit elements' states, marking the moment of switching in by a plus sign, and the moment of switching out by a minus sign [15, 16]. An example of such a table is shown in Fig. 6 (in its number of inputs and outputs, this table corresponds to Fig. 5).

If the feedback path can meaningfully be sundered, a sequential circuit can then be transformed to a nonsequential one (Fig, 5,b), and its operating conditions can be described in the form of the so-called nonsequential equivalent (Cf. the table in Fig. 7).*

This circumstance permits the reduction of the problem of minimizing Boolean functions for sequential circuits to the same problem for nonsequential circuits, and allows the methods presented earlier to be used. However, the presence of a definite temporal sequence in which the circuit's input and output states follow each other, allows, in this case, a significant reduction in, or complete avoiding of, the sorting operation.

---

* This concept was introduced by P. P. Parkhomenko.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$-A$	$+A$																			
$-Z_1$		$+Z_1$						$-Z_1$					$+Z_1$			$-Z_1$				$+Z_1$
$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							$-Z_3$		
$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
$-Z_5$				$+Z_5$								$-Z_5$								

etc.

Fig. 6.

Z_5	Z_4	Z_3	Z_2	Z_1	A	Z_5	Z_4	Z_3	Z_2	Z_1
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	1
0	0	0	0	1	1	0	0	0	1	1
0	0	0	1	1	1	1	0	0	1	1
0	0	1	0	0	1	0	0	1	0	1
0	0	1	0	1	1	0	0	1	1	1
0	0	1	1	1	1	0	1	1	1	1
0	1	0	0	0	1	0	0	0	0	0
0	1	1	0	0	1	0	1	0	0	0
0	1	1	1	0	1	0	1	1	0	0
0	1	1	1	1	1	0	1	1	1	0
1	0	0	0	0	1	1	0	1	0	0
1	0	0	1	1	1	1	0	1	1	1
1	0	1	0	0	1	0	0	1	0	1
1	0	1	1	1	1	1	1	1	1	1
1	1	0	0	0	1	1	0	0	0	0
1	1	0	1	0	1	1	0	0	0	0
1	1	0	1	1	1	1	1	0	1	0
1	1	1	1	1	1	1	1	0	1	1

Fig. 7.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-A$	A																			
$2'$	$-Z_2$					$+Z_2$		$-Z_2$				$+Z_2$							$-Z_2$		
2^2	$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
2^3	$-Z_3$				$+Z_3$								$-Z_3$								
$Z_1=1$		1	1	1	9	11	15						3	3	3					1	1
$Z_1=0$	0							13	13	13	9	11				7	7	7	5		

a

	0	1	2	3	4	5	6	7	8	9	10	11	12.	13	14	15	16	17	18	19	20
2^0	$-Z_1$		$+Z_1$						$-Z_1$					$+Z_1$			$-Z_1$				$+Z_1$
$Z_2=1$			1	1	1	1	1	1						1	1	1					1
$Z_2=0$	0	0							0	0	0	0	0				0	0	0	0	

b

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
$2'$	$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
2^2	$-Z_5$				$+Z_5$								$-Z_5$								
$Z_3=1$					5	5					4	4	0	0	1	3	3				
$Z_3=0$	0	0	0	1			7	7	7	6								2	2	0	0

c

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							$-Z_3$		
$2'$	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
$Z_4=1$						3	3	2	2						3	3	3	1			
$Z_4=0$	0	0	0	2	2					0	0	1	1	1					0	0	0

d

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
$2'$	$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							$-Z_3$		
$Z_5=1$				1	1	3	3	1	1	0	0										
Z_5 0	0	0	0									2	2	2	3	3	3	2	0	0	0

e

Fig. 8.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-A$	$+A$																			
2^1	$-Z_3$					$+Z_3$		$-Z_3$				$+Z$							$-Z_3$		
2^2	$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
2^3	$-Z_5$				$+Z_5$								$-Z_5$								
2^4	$-Z_2$		$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$				
$Z_1=1$		1	17	17	25	27	31						3	19	19					1	17
$Z_1=0$	0							29	13	13	9	11				23	7	7	5		

a

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
2^1	$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
2^2	$-Z_5$				$+Z_5$								$-Z_5$								
2^3	$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							$-Z_3$		
$Z_2=1$					5	13					4	12	8	8	9	11	11				
$Z_2=0$	0	0	0	1			15	7	7	6								10	2	0	0

b

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							Z_3		
2^1	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
2^2	$-Z_4$						$+Z_4$				$-Z_4$					$+Z_4$				$-Z_4$	
$Z_4=1$						3	7	6	6						3	7	7	5			
$Z_4=0$	0	0	0	2	2					4	0	1	1	1					4	0	0

c

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2^0	$-Z_2$			$+Z_2$						$-Z_2$					$+Z_2$			$-Z_2$			
2^1	$-Z_3$					$+Z_3$		$-Z_3$				$+Z_3$							$-Z_3$		
2^2	$-Z_5$				$+Z_5$								$-Z_5$								
$Z_5=1$				1	5	7	7	5	5	4	4										
$Z_5=0$	0	0	0									6	2	2	3	3	3	2	0	0	0

d

Fig. 9.

As was shown in [16, 18], in the Boolean functions characterizing each of the outputs of a sequential circuit, there must be contained, as a minimum, the terms characterizing the elements which change state when the circuit of a given output must be closed (in a sequential switching table, at the cycles designed for the switching in of the elements) and when the circuit must be opened (at the cycles designed for the switching off of the elements). If the switching table which contains the sequence of switching on and off of these elements is realizable, then the Boolean function for the given output can not contain terms which correspond to any other element. If the switching table for the aforementioned elements is not realizable, it is then necessary to add to it, from the original switching table, a minimal number of sequences of these elements which would make it realizable. This operation still does not give a minimal form of the Boolean function, but does permit the immediate discarding of the letters which will not enter into all its terms which, certainly, significantly simplifies the minimization problem.*

If we carry out this operation on the table of Fig. 6, then, for the individual elements, we obtain the tables given in Fig. 8,a,b,c,d and e. As may be seen from Fig. 8, the switching table obtained is realizable only for element Z_2. Since one sequence of switching of element Z_1 occurs in this table, the minimal Boolean function for Z_2 will then have the form:

$$F_{(Z_2)\min} = z_1.$$

The remaining switching tables can be made realizable if sequences of actions of the very elements which comprise them are added to them (Fig. 9).

The method just considered thus allows one to decompose an ordinary switching table for a sequential circuit into a number of sub-tables with fewer numbers of elements. The minimization problem is thereby significantly simplified.

However, the special features of sequential circuits allow an even greater simplification of the minimization problem to be made.

As was shown in [16], the transition from a sequential switching table to the Boolean function realizing condition f_i can be carried out by means of the formula

$$F_{i(Z)} = f_{\text{op}} + zF_i^{\prime\prime}, \qquad (23)$$

where f_{op} is a sum of the terms which correspond to the combinations of circuit elements states in the first cycle of each of the switched-in periods of element Z, but without the state of this element itself being taken into account, and F_i'' is the sum of the terms which correspond to combinations of circuit element states in the other cycles wherein the Z circuit must be closed, again with the state of element Z being left out of account.

We shall call f_{op} the operated state, and zF_i'' the switched-in state.

Theorem 3. A particular minimal form of the Boolean function which realizes the switched-in states accurately with the possible exception of the letter $\underline{z}$ may be obtained by choosing such a minimal number and such factors which, in their totality, would overlap all the cycles of the switched-in state, exactly coinciding with the succeeding cycles following each of the switched-in periods of element Z.

Indeed, all the terms forming zF_i'' have the factor $\underline{z}$ which equals one only for f_i. Therefore, this factor in conjunction with any other term φ' which satisfies the conditions given above, gives a term φ. It must thus be proved that this term will be minimal, or differ from the minimal by the letter $\underline{z}$.

If the element Z, in the switching table, has one operated period, i.e., is switched in and switched out once, it then suffices, for obtaining overlap, to choose a term φ' which contains one letter. It is obvious that an exact coincidence with the following cycles of F_i'' can be guaranteed if, as one of the φ', a letter is selected which corresponds to an element which changes its state after the circuit of element Z is opened. Since an exact coincidence with the beginning of F_i'' is not required, it is then always possible to find such letters which, with a definite overlap, would cover the remaining cycles of F_i'' without redundant repetitions.

*An analogous operation can also be carried out on switching tables for nonsequential circuits: in them can be stricken on individual columns, thus decreasing the number of letters involved in the sorting, until the point is reached when at least one f_i coinciding with an f_0 is obtained.

	z_2	z_3	z_5	
	0	0	1	f_1
	1	0	0	f_1
	1	0	1	f_1
	1	1	1	f_1
	0	0	0	f_0
	0	1	0	f_0
	0	1	1	f_0
	1	1	0	f_0

a

	A	z_1	z_3	z_4	z_5	
1	1	0	0	0	0	f_1
2	1	0	1	0	0	f_1
3	1	1	0	0	0	f_1
4	1	1	0	0	1	f_1
5	1	1	1	0	0	f_1
6	1	1	1	0	1	f_1
7	1	1	1	1	1	f_1
	0	0	0	0	0	f_0
	1	0	0	0	1	f_0
	1	0	0	1	0	f_0
	1	0	1	0	1	f_0
	1	0	1	1	0	f_0
	1	1	0	1	1	f_0
	1	1	1	1	0	f_0

b

Fig. 10.

f_1	$A = z_1\bar{z}_4$	$B = z_1z_3z_5$	$C = z_3\bar{z}_4z_5$	$D = z_1\bar{z}_3\bar{z}_5$	$E = a\bar{z}_4\bar{z}_5$
1					+
2			+		+
3	+			+	+
4	+				
5	+		+		+
6	+	+			
7		+			

Fig. 11.

With these conditions:

a) none of the terms φ' can be eliminated, since this calls for going counter to the condition of overlapping all cycles of F_1'', i.e., at least one of the conditions f_1 entering into F_1'' would be unimplemented; thus, the number of terms in this case is minimal, i.e., corresponds to one of the particular minimal forms for zF_1''.

b) the terms φ all turn out to consist of two letters, one of which will be the letter $\underline{z}$, i.e., will in fact differ from the minimal by no more than $\underline{z}$.

For example, we consider the table for element Z_5 in the example given above (Fig. 9,d).

The term f_{op}, in correspondence with the element states at cycle 3, will equal

$$f_{op} = z_2\bar{z}_3.$$

Switching off of element Z_5 occurs after operation of element Z_3. Therefore, to provide exact coincidence with the cycles following the switched-in state, it is necessary to choose term $\bar{z}_3$ as one of the φ'. To overlap all the switched-in state cycles, it is necessary to add $\varphi' = z_2$ to this. Thus, we shall have .

$$F_{(Z_5)} = z_2\bar{z}_3 + z_5(z_2 + \bar{z}_3).$$

If we set up the enumeration of conditions f_1 and f_0 (Fig. 10,a) for the table of Fig. 9,d and then determine the minimal implicants by means of the first algorithm, we shall obtain the same results.

Now, let element Z have several operated periods. In theory, it is possible to overlap all the switched-in states in the same manner as before. It may turn out, however, that a term φ', consisting of one letter and providing exact coincidence with the succeeding cycles of the switched-in state may, in different operated periods of element Z, give rise to contradictory conditions. Thus, for example, in the example considered above for element Z_1 (Fig. 9,a), the term which provides coincidence with the succeeding cycles in the first period is the term $\varphi_1' = z_3$ and, in the second period, the term $\varphi_2' = z_4$. However, in this same period, element z_3 is switched off at the limits of F_1'' and, therefore, the term z_1z_3 will disrupt the given sequence of actions of Z_1, retaining the value $zF_1'' = 1$ even at the limits of F_1''.

This is indicative of the fact that, in this case, the term φ', which is one letter long, has a length less that minimal. For such terms we seek a set of letters, of minimal content, which would liquidate the contradiction noted.

The expressions thus obtained, realizing the switched-in state, will have in each term a minimal number of letters necessary for the conditions of realizing the switched-in state, and a minimal number of terms themselves, and will also have the letter $\underline{z}$ making them up, i.e., will differ from the possible minimal terms by not more than $\underline{z}$.

Thus, for example, for Z_1', in correspondence with the switching table of Fig. 9,a, we will have:

for the first operated period

$$f_{op} = a\bar{z}_3\bar{z}_4\bar{z}_5 \text{ and } zF^{*}_{i,1} = z_1(z_3 + \bar{z}_4);$$

and for the second operated period*

$$f_{op} = az_3\bar{z}_4\bar{z}_5 \text{ and } zF^{*}_{i,2} = z_1\bar{z}_4.$$

Since, as was stated, the term $\varphi' = z_3$ contradicts the succeeding cycles of the switched-in state for the second operated period, we add to it an additional letter which would differentiate the first operated period from the second. Such a letter, as may be seen from the table of Fig. 9,a, is z_5. Thus,

$$\begin{aligned} F_{(Z_1)} &= a\bar{z}_3\bar{z}_4\bar{z}_5 + az_3\bar{z}_4\bar{z}_5 + z_1(z_5z_3 + \bar{z}_4 + \bar{z}_4) = \\ &= a\bar{z}_4\bar{z}_5 + z_1(\bar{z}_4 + z_3z_5). \end{aligned} \tag{24}$$

If, in constructing the table of conditions for f_1 and f_0 for Z_1, we employ the first algorithm considered (Fig. 10,b), we obtain a general minimal form of the Boolean function in the form

$$F_{(Z_1)} = z_1\bar{z}_4 + z_1z_3z_5 + z_3\bar{z}_4\bar{z}_5 + z_1\bar{z}_3\bar{z}_5 + a\bar{z}_4\bar{z}_5. \tag{25}$$

By using McCluskey's method (cf. the table in Fig. 11), we obtain

$$\Phi = E(C \vee E)A \vee D \vee E)A(A \vee C \vee E)(A \vee B)B = ABE.$$

Thus, the general minimal form has the form:

$$F_{(Z_1)} = z_1\bar{z}_4 + z_1z_3z_5 + a\bar{z}_4\bar{z}_5,$$

i.e., coincides completely with the result obtained above.

To obtain the general minimal form, it is necessary to take, for each of the outputs, all the variants of the switching table to be realized which have a minimal number of elements and, for each of these tables, to take all the variants of overlapping the cycles of F_1''.

We consider one more example of the sequential-nonsequential circuit. Let the switching condition of element Z be the application of stimuli at inputs A, B, C and D in the following order: +A, +B, −A, +C, −B, +A, and +D. Any disruption of this sequence must cause element Y to switch in. For element Z the operating conditions are sequential, since it must be switched in after a definite unambiguously given temporal sequency of stimuli applications to inputs. For element Y these conditions are sequential-nonsequential, since this element must not operate for the sequence of stimuli given for Z, and must operate for all other sequences, whose temporal orders are not given.

If we consider each cycle of the correct sequence of application of stimuli, and after these if we enumerate all the possible incorrect stimuli, for which element Y must operate, we obtain the table of conditions in f_1 and f_0 which is given in Fig. 12 (the unprimed cycle numbers denote cycles of the given sequence while the primed cycle numbers denote cycles in which incorrect stimuli can occur).

The complete switching table, with weighed states, for Z and Y is given in Fig. 13,a. For element Y, this table is unrealizable. An analysis of the conditions in it [15, 16] shows that three additional elements must be included in it in order to make it realizable. The switching table with these additional elements included is given in Fig. 13,b.

Element Z is switched in after a stimulus is applied to input D. However, to obtain a realizable switching table, one such element turns out to be insufficient. For this it is necessary to introduce a sequence of elements A and C in the table (Fig. 13,c).

* The third period (cycles 19 and 20) is a repetition of the first.

Cycle	State Weight	A (2^0)	B (2^1)	C (2^2)	D (2^3)	Y
0	0	0	0	0	0	0
0′	2	0	1	0	0	1
	4	0	0	1	0	1
	8	0	0	0	1	1
1	1	1	0	0	0	0
1′	0	0	0	0	0	1
	5	1	0	1	0	1
	9	1	0	0	1	1
2	3	1	1	0	0	0
2″	1	1	0	0	0	1
	7	1	1	1	0	1
	11	1	1	0	1	1
3	2	0	1	0	0	0
3′	3	1	1	0	0	1
	0	0	0	0	0	1
	10	0	1	0	1	1
4	6	0	1	1	0	0
	7	1	1	1	0	1
	2	0	1	0	0	1
	14	0	1	1	1	1
5	4	0	0	1	0	0
5′	6	0	1	1	0	1
	0	0	0	0	0	1
	12	0	0	1	1	1
6	5	1	0	1	0	0
6′	4	0	0	1	0	1
	7	1	1	1	0	1
	1	1	0	0	0	1
7	13	1	0	1	1	0

Fig. 12.

	X_3 (2^6)	X_2 (2^5)	X_1 (2^4)	D (2^3)	C (2^2)	B (2^1)	A (2^0)	State Weight
1	0	0	0	0	0	1	0	2
2	0	0	0	0	1	0	0	4
3	0	0	0	1	0	0	0	8
4	0	0	1	0	0	0	0	16
5	0	0	1	0	1	0	1	21
6	0	0	1	1	0	0	1	25
7	0	1	0	0	0	0	0	32
8	0	1	0	0	0	1	1	35
9	0	1	0	1	0	1	0	42
10	0	1	1	0	0	0	1	49
11	0	1	1	0	1	1	1	55
12	0	1	1	1	0	1	1	59
13	1	0	0	0	0	0	0	64
14	1	0	0	0	1	1	0	70
15	1	0	0	1	1	0	0	76
16	1	0	1	0	0	0	1	81
17	1	0	1	0	1	0	0	84
18	1	0	1	0	1	1	1	87
19	1	1	0	0	0	1	0	98
20	1	1	0	0	1	1	1	103
21	1	1	0	1	1	0	0	110
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	1	1
	0	0	1	0	0	0	1	17
	0	0	1	0	0	1	1	19
	0	1	0	0	0	1	0	34
	0	1	0	0	1	1	0	38
	0	1	1	0	0	1	0	50
	0	1	1	0	0	1	1	51
	1	0	0	0	1	0	0	68
	1	0	0	0	1	0	1	69
	1	0	1	0	1	0	1	85
	1	0	1	1	1	0	1	93
	1	1	0	0	1	0	0	100
	1	1	0	0	1	1	0	102

Fig. 14.

	1	4	8	16	21	25	32	35	42	49	55	59	64	70	76	81	84	87	98	100	110
$A = d\bar{a}$			+						+						+						+
$B = db$									+			+									
$C = d\bar{c}$			+			+			+			+									
$D = abc$											+							+		+	
$F_1 = \bar{x}_1 d$			+						+						+						+
$F_2 = \bar{x}_1 ba$								+												+	
$G_1 = x_1\bar{b}\bar{a}$				+													+				
$G_2 = x_1 c\bar{a}$																	+				
$G_3 = x_1 cb$											+							+			
$I_1 = \bar{x}_2 b\bar{a}$	+													+							
$I_2 = \bar{x}_2 cb$														+				+			
$I_3 = \bar{x}_2 x_1\bar{a}$				+													+				
$I_4 = \bar{x}_2\bar{x}_1 b$	+													+							
$K_1 = x_2\bar{b}a$										+											
$K_2 = x_2 ca$											+									+	
$K_3 = x_2 c\bar{b}$							+			+											
$K_4 = x_2\bar{x}_1 a$								+												+	
$K_5 = x_2 x_1\bar{b}$										+											
$K_6 = x_2 x_1 c$											+										
$K_7 = x_2 d$									+			+									+
$L_1 = \bar{x}_3 d$			+			+			+			+									
$L_2 = \bar{x}_3 ca$					+						+										
$L_3 = \bar{x}_3 c\bar{b}$			+		+																
$L_4 = \bar{x}_3 x_1 c$					+						+										
$L_5 = \bar{x}_3 x_2\bar{b}$							+			+											
$L_6 = \bar{x}_3\bar{x}_2 c$		+			+																
$M_1 = x_3\bar{c}$													+			+			+		
$M_2 = x_3 ba$																		+		+	
$M_3 = x_3 x_1\bar{a}$																	+				
$M_4 = x_3 x_1 b$																		+			
$M_5 = x_3 x_2 a$																				+	
$M_6 = x_3\bar{x}_2 b$														+				+			

Fig. 15.

	0	0′	1	1′	2	2′	3	3′	4	4′	5	5′	6	6′	7
2^0	$-A$		$+A$				$-A$						$+A$		
2^1	$-B$				$+B$						$-B$				
2^2	$-C$								$+C$						
2^3	$-D$														$+D$
$Z=0$	0	0	1	1	3	3	2	2	6	6	4	4	3	3	
$Z=1$															13
$Y=0$	0	0	1	1	3	3	2	2	6	6	4	4	3	3	13
$Z=0$ $Y=0$		2 4 8		0 5 9		1 7 11		3 0 10		7 2 14		6 0 12		4 7 1	

a

	0	0′	1	1′	1″	2	2′	2″	3	3′	3″	4	4′	4″	5	5′	5″	6	6′	6″	7
2^0	$-A$		$+A$						$-A$									$+A$			
2^1	$-B$					$+B$									$-B$						
2^2	$-C$											$+C$									
2^3	$-D$																				$+D$
2^4	$-X_1$			$+X_1$						$-X_1$									$+X_1$		
2^5	$-X_2$						$+X_2$									$-X_2$					
2^6	$-X_3$												$+X_3$								
$Z=0$	0	0	1	17	17	19	51	51	50	34	34	38	102	102	100	68	68	69	85	85	
$Z=1$																					93
$Y=0$	0	0	1	17	17	19	51	51	50	34	34	38	102	102	100	68	68	69	85	85	93
$Z=0$ $Y=1$		2 4 8			16 21 25			49 55 59			35 32 42			103 98 110			70 64 76			84 87 81	

b

	0	0′	1	1′	2	2′	3	3′	4	4′	5	5′	6	6′	7
2^0	$-A$		$+A$				$-A$						$+A$		
2^1	$-C$								$+C$						
2^2	$-D$														$+D$
$Z=1$															7
$Z=0$	0	0 2 4	1	1 0 3 5	1	1 3 5	0	0 1 4	2	2 3 0	2	2 0 6	3	3 2 1	

c

Fig. 13.

For element Y we apply the algorithm flowing from Theorem 1, since its operating conditions are non-sequential. The table of conditions f_1 and f_0 for it are given in Fig. 14. By carrying out analysis of the minimal terms, we obtain the general minimal form in the form

$$\begin{aligned} F_{(Y)} = {} & d\bar{a} + db + d\bar{c} + abc + \bar{x}_1 d + \bar{x}_1 ba + x_1 \bar{b}\bar{a} + \\ & + x_1 c\bar{a} + x_1 cb + \bar{x}_2 b\bar{a} + \bar{x}_2 cb + \bar{x}_2 x_1 \bar{a} + \bar{x}_2 \bar{x}_1 b + \\ & + x_2 \bar{b} a + x_2 ca + x_2 \bar{c}\bar{b} + x_2 \bar{x}_1 a + x_2 x_1 \bar{b} + x_2 x_1 c + \\ & + x_2 d + \bar{x}_3 d + \bar{x}_3 ca + \bar{x}_3 c\bar{b} + \bar{x}_3 x_1 c + \bar{x}_3 x_2 \bar{b} + \\ & + \bar{x}_3 \bar{x}_2 c + x_3 \bar{c} + x_3 ba + x_3 x_1 \bar{a} + x_3 x_1 b + x_3 x_2 a + x_3 \bar{x}_2 b. \end{aligned}$$

The table for realizing the conditions f_1 according to McCluskey are given in Fig. 15. The formula for overlapping the conditions f_1 deriving from it will have the form:

$$\begin{aligned} \Phi = {} & (I_1 \vee I_4) L_6 (A \vee C \vee F_1 \vee L_1 \vee L_3)(C_1 \vee I_3)(L_2 \vee L_3 \vee \\ & \vee L_4 \vee L_6)(C \vee L_1)(K_3 \vee L_5)(F_2 \vee K_4)(A \vee B \vee C \vee F_1 \vee \\ & \vee K_7 \vee L_1)(K_1 \vee K_3 \vee K_5 \vee L_5)(D \vee G_3 \vee K_2 \vee K_6 \vee L_2 (L_4) \times \\ & \times (B \vee C \vee K_7 \vee L_1) M_1 (I_1 \vee I_2 \vee I_4 \vee M_6)(A \vee F_1) M_1 \times \\ & \times (G_1 \vee G_2 \vee I_3 \vee M_3)(D \vee G_3 \vee I_2 \vee M_2 \vee M_4 \vee M_6) M_1 \times \\ & \times (D \vee F_2 \vee K_2 \vee K_4 \vee M_2 \vee M_5)(A \vee F_1 \vee K_7). \end{aligned}$$

By transforming this formula, we get

$$\begin{aligned} \Phi = {} & L_6 M_1 (A \vee F_1)(I_1 \vee I_4)(G_1 \vee I_3)(C \vee L_1)(K_3 \vee L_5) \times \\ & \times (F_2 \vee K_4)[D \vee G_3 \vee (K_2 \vee K_6 \vee L_2 \vee L_4)(I_2 \vee M_2 \vee M_4 \vee M_6)]. \end{aligned}$$

We can obtain all the particular minimal forms from this formula. One of them, containing the least number of terms, is obtained if we take the conjunction consisting of the first terms of the bracketed expressions in the logical formula, i.e., $L_6M_1AI_1G_1CK_3F_2D$. The minimal form of the Boolean function which corresponds to this will have the form:

$$\begin{aligned} F_{(Y)} = {} & \bar{x}_3 \bar{x}_2 c + x_3 \bar{c} + d\bar{a} + \bar{x}_2 b\bar{a} + x_1 \bar{b}\bar{a} + d\bar{c} + \\ & + x_2 \bar{c}\bar{b} + \bar{x}_1 ba + abc = d(\bar{a} + \bar{c}) + ab(c + \bar{x}_1) + \\ & + \bar{c}(x_3 + x_2 \bar{b}) + x_1 \bar{b}\bar{a} + \bar{x}_3 \bar{x}_2 c. \end{aligned}$$

Thus, the minimization methods considered above allow a rapid determination to be made of the minimal forms of Boolean functions for quite complicated conditions.

Received August 12, 1958

LITERATURE CITED

[1] C. Shannon, A Symbolic Analysis of Relay and Switching Circuits, Trans. of American Institute of Electr. Engineers, 1938.

[2] Synthesis of electronic computing and control circuits. The Annals of the Computation laboratory of Harvard University, vol. 27, 1951.*

* Russian translation, published by IL in 1954.

[3] W. V. Quine, The problem of simplifying truth function. The American Mathematical Monthly, vol. 59, No. 8, 1952.

[4] W. V. Quine, Two theorems about truth functions. Boletin de la sociedad Matematica Mexicana, vol. 10, No. 1-2, 1953.

[5] R. H. Urbano and R. K. Mueller, A topological method for the determination of minimal forms of Boolean function. Trans. IRE, vol. Ec-5, No. 3, Sept., 1956.

[6] J. McCluskey, Minimization of Boolean functions. B.S.T.J., vol. XXXV, No. 6, 1956.

[7] N. B. Acred, Control Circuit Design. Electronic Engineering, vol. 29, Dec., 1957.

[8] B. Harris, An algorithm for determining minimal representation of a logic function. Trans. IRE, vol. Ec-6, No. 1, 1957.

[9] M. Marcus, Minimization of the partially developed transfer tree. Trans. IRE, vol. Ec-6, No. 2, 1957.

[10] V. N. Roginskii, "Taking unused states into account in the synthesis of switching circuits," [in Russian], Automation and Remote Control (USSR) 15, 3 (1954).

[11] Ya. I. Mekler, "Sufficient constituents for the operation of switching circuits," [in Russian] Dokl. AN SSSR 117, 14 (1957).

[12] A. N. Yurasov, "On the question of setting up structural formulae for sequential circuits," [in Russian]. In the collection, Devices and Elements of Automation and Remote Control [in Russian], Izd. Mosk. Mekhanich. In-ta (1952).

[13] A. Svoboda, "Certain applications of contact circuits," [in Russian]. In collection, Subjects of Papers at the Moscow All-Union Conference on the Theory of Relay Devices [in Russian]. Izdatelstvo AN SSSR (1957).

[14] K. K. Voishvillo, "A method for simplifying the form of truth functions," [in Russian]. Nauchnye doklady vysshei shkoly. Filosofskie nauki 2 (1958).

[15] M. A. Gavrilov, Theory of Relay-Contact Circuits [in Russian] (page 521). Izdatelstvo AN SSSR (1950).

[16] M. A. Gavrilov, "Basic formulae of relay circuit synthesis," [in Russian], Automation and Remote Control (USSR) 15, 6 (1954).

[17] R. McNaughton and B. Mitchell, The Minimality of Rectifier Nets with Multiple Outputs Incompletely Specified. Frankl. Inst., vol. 264, No. 6 (1957).

[18] M. A. Gavrilov, Structural Theory of Switching Circuits (Lectures). [in Russian], Izdatelstvo VZÉI (1956).

POSSIBILITIES OF INCREASING NOISE STABILITY BASED ON THE USE OF A PRIORI PARAMETER PROBABILITIES

V. A. Kashirin and G. A. Shastova
(Moscow)

The noise stability of an ideal receiver is considered, the first distribution law of the a priori parameter value probabilities being taken into account. An analysis of noise stability is carried out for the case of a nonlinear parameter transformation prior to modulation. The optimal nonlinear transformation of parameters prior to transmission is found, as well as the gain in noise stability obtained with this transformation. Both the mean square and the information criteria of optimality are used.

It is shown in the theory of potential noise stability [1] that the receiver which identifies its received signals with their most probable parameter values is ideal for weak fluctuating noise. The noise stability of such a receiver was investigated in detail by V. A. Kotel'nikov for the case of equiprobable transmission of all parameter values in some range. In practice, the parameter probability distribution law is frequently unknown. In this case, the receiver is generally chosen on the basis that all parameter values have equal probability of being transmitted. It is advantageous to consider the noise stability of an ideal receiver, designed for a known first probability distribution law of the parameters, this law differing from a uniform distribution, and to compare its noise stability with the noise stability of a receiver constructed for the case of an unknown first distribution law. This allows us to judge what a knowledge of the distribution law can give us in terms of increasing the noise stability of reception.

Noise Stability of an Ideal Kotel'nikov Receiver for a Nonuniform Parameter Probability Distribution

Let the parameter λ, varying with a certain probability distribution $P(\lambda)$ within the limits of ± 1, be transmitted by means of a signal $A(\lambda, t)$ along a channel with weak fluctuating noise of intensity σ, in $\mathrm{cps}^{-1/2}$. By using the methods of potential noise stability theory, we find the probability that, if the signal $x(t)$ is received, the parameter λ lies within the limits $\lambda' < \lambda < \lambda' + d\lambda$. This probability equals

$$P[\lambda' < \lambda < \lambda' + d\lambda] = \frac{\exp\left\{-\frac{T}{\sigma^2}\overline{[x(t) - A(\lambda', t)]^2} + \ln P(\lambda')\right\} d\lambda}{\int\limits_{-1}^{1}\exp\left\{-\frac{T}{\sigma^2}\overline{[x(t) - A(\lambda, t)]^2} + \ln P(\lambda)\right\} d\lambda} = P_x(\lambda')d\lambda.$$

Here, T is the interval during which the parameter value is taken to be unchanged and

$$\overline{[x(t) - A(\lambda, t)]^2} = \frac{1}{T}\int\limits_0^T [x(t) - A(\lambda, t)]^2\, dt.$$

The function $P_x(\lambda')$ characterizes the probability of various errors in the reproduction of the parameter when x(t) is received. It can be written in the form

$$P_x(\lambda') = k_x \exp\left\{-\frac{T}{\sigma^2}\overline{[x(t)-A(\lambda', t)]^2} + \ln P(\lambda')\right\}.$$

Here, k_x is a coefficient which depends on $\underline{x}$ but not on λ' or $\underline{t}$. For a given x(t), the most probable value of the parameter λ_{xn} must minimize the following expression

$$\left\{\frac{T}{\sigma^2}\overline{[x(t)-A(\lambda', t)]^2} - \ln P(\lambda')\right\}_{\lambda'=\lambda_{xn}} = \min. \tag{1}$$

With weak noise the errors on the receiving end are small and it is therefore possible to consider only the probabilities of reproducing parameter values close to the most probable value. This allows us to discard the higher terms in the expansion of function $A(\lambda, t)$ in the neighborhood of λ_{xn}, i.e., to consider that

$$A(\lambda, t) = A(\lambda_{xn}, t) + A'_\lambda(\lambda_{xn}, t)(\lambda - \lambda_{xn}). \tag{2}$$

Here,

$$A'_\lambda(\lambda_{xn}, t) = \left[\frac{\partial A(\lambda_{xn}, t)}{\partial \lambda}\right]_{\lambda=\lambda_{xn}}.$$

By using (2), as well as the expression

$$\int_{-\infty}^{\infty} P_x(\lambda')\, d\lambda' = 1, \tag{3}$$

we can define k_x and, consequently, the function $P_x(\lambda')$.

The magnitude of the mean square error equals

$$\delta^2 = \int_{-\infty}^{\infty} (\lambda' - \lambda_{xn})^2 P(\lambda')\, d\lambda'. \tag{4}$$

It should be noticed that, when (3) and (4) are integrated between infinite limits, an error is admitted but, for weak noise, $P_x(\lambda')$ falls so rapidly with increasing $(\lambda'-\lambda_{xn})$ that the integration between infinite limits does not entail any essential error. Solving the problem posed in its general form gives rise to mathematical difficulties. We consider the particular case wherein the parameter is distributed normally, symmetrically about zero:

$$P(\lambda') = qe^{-\frac{\lambda'^2}{2\lambda_0^2}} \quad (1 \geqslant \lambda \geqslant -1). \tag{5}$$

The probability that, if $\underline{x}$ is received, parameter λ' was transmitted will equal

$$P_x(\lambda') = k'_x \exp\left\{-\frac{T}{\sigma^2}\overline{[A'^2_\lambda(\lambda_{xn}, t)(\lambda'-\lambda_{xn})^2]} - \frac{\lambda_{xn}}{\lambda_0^2}(\lambda'-\lambda_{xn}) - \frac{\lambda'^2}{2\lambda_0^2}\right\}. \tag{6}$$

From condition (1) we easily obtain

$$\frac{2T}{\sigma^2}\overline{[x(t)-A(\lambda_{xn},\ t)]\,A'_\lambda(\lambda_{xn},\ t)}=\frac{\lambda_{xn}}{\lambda_0^2}.$$

By substituting (6), in (3) and determining k_x', we find the error probability function:

$$P_x(\lambda')=e^N\sqrt{\frac{p}{\pi}}\,e^{-(\lambda'-\lambda_{xn})^2p-N},$$

where

$$N=\frac{\lambda_{xn}^2}{2\lambda_0^2},\quad p=\frac{T}{\sigma^2}\overline{A_\lambda'^2(\lambda_{xn},\ t)}+\frac{1}{2\lambda_0^2}.$$

We find, from formula (4), the magnitude of the mean square error:

$$\delta_N^2=\frac{\sigma^2}{8T\overline{A_\lambda'^2(\lambda_{xn},\ t)}+4\frac{\sigma^2}{\lambda_0^2}}=\frac{\delta_0^2}{1+4\frac{\delta_0^2}{\lambda_0^2}}.$$

Here, δ_0 is the mean square error given by the ideal receiver when the signal statistics are not taken into account:

$$\delta_0^2=\frac{\sigma^2}{8T\,\overline{A_\lambda'^2(\lambda_{xn},\ t)}}. \tag{7}$$

When the statistics are taken into account, the error is decreased by a factor of B_N:

$$B_N=\frac{\delta_0^2}{\delta_N^2}=1+4\frac{\delta_0^2}{\lambda_0^2}. \tag{8}$$

It follows from (8) that an appreciable gain in noise stability is obtained only for $\lambda_0 < 4\delta_0$, i.e., for very small λ_0, since δ_0 is small in the case of weak noise. From the practical point of view, the case $\lambda_0 < 4\delta_0$ is rarely realized. Ordinarily, $\lambda_0 >> 4\delta_0$, and then $\delta_N \approx \delta_0$.

Thus, in the majority of practical cases, the ideal Kotel'nikov receiver provides no increase in noise stability when the parameter's first distribution law is taken into account. An essential disadvantage of a receiver which takes the first distribution law into account is that its operating mode depends on the ratio of the spectral density of the noise to the parameter's dispersion. For some definite ratio, σ^2/λ_0^2, this receiver will provide minimum error, but if this ratio is changed, the receiver will not provide this minimum. In particular, if the receiver is designed for the reception of signals with a definite noise level, it will then give an error when the transmission is noiseless, this error being of the same order of magnitude as the gain which the receiver provides when the statistics are taken into account as compared to a receiver which does not take the statistics into account. We shall illustrate this with one example.

Let the signal be transmitted by means of amplitude modulated dc, i.e., $A(\lambda,\ t) = U_m\lambda$, and let the parameter obey a normal law, i.e.,

$$P(\lambda)=qe^{-\frac{\lambda^2}{2\lambda_0^2}}.$$

The ideal receiver reproduces the most probable parameter value which satisfies condition (1). By computing the defivative of $P(\lambda)$ with respect to λ for $\lambda = \lambda_{xn}$, we obtain the equation defining λ_{xn}:

$$-\frac{2TU_m}{\sigma^2}\left[x(t) - U_m\lambda_{xn}\right] - \frac{\lambda_{xn}}{\lambda_0^2} = 0.$$

From whence,

$$\lambda_{xn} = \frac{1}{1 - \frac{1}{\lambda_0^2}\frac{\sigma^2}{2TU_m^2}}\frac{x(t)}{U_m}.$$

By taking into account that

$$\frac{\sigma^2}{8TU_m^2} = \delta_0^2, \tag{9}$$

we obtain the following expression for the most probable value of the parameter λ for the noise level defined by equation (9):

$$\lambda_{xn} = \frac{1}{1 - 4\frac{\delta_0^2}{\lambda_0^2}}\frac{x(t)}{U_m}.$$

If a signal without noise acts on this receiver, i.e., $x(t) = U_m\lambda_{trans}$, it then turns out that the received value, λ_{rec}, will differ from the transmitted value, λ_{trans}:

$$\lambda_{rec} = \frac{1}{1 - 4\frac{\delta_0^2}{\lambda_0^2}}\lambda_{trans}.$$

For $\lambda_0 \gg 4\delta_0$, the error of reproduction will equal the ratio $(4\delta_0/\lambda_0)^2$, i.e., the gain given by the receiver when the statistics are taken into account.

The analysis given shows that, in the case when the dispersion of the transmitted random parameter is many times larger than the admissible magnitude of the mean square error, the use of the knowledge of the first probability distribution law in the ideal Kotel'nikov receiver provides a negligible gain in noise stability and gives rise to an additional error when the noise level changes.

Nonlinear Parameter Transformation

We now consider the influence of the method of modulation on noise stability from the point of view of using signal statistics. When a parameter λ is transmitted through a channel with fluctuating noise to an ideal receiver, the mean square error is defined, as is well-known [1], by expression (7).

The mean square error, averaged over all values of parameter will be

$$\bar{\delta}^2 = \frac{\sigma^2}{8}\int_{-1}^{1}\frac{P(\lambda)}{\int_0^T\left[\frac{\partial A(\lambda, t)}{\partial\lambda}\right]^2 dt}d\lambda. \tag{10}$$

It follows from (10) that the mean square error depends on both the function P(λ) and the function A(λ, t). Obviously, for each given function P(λ) one can find the optimal function A(λ, t) which would minimize the mean square error. In the majority of modulation systems, the function A(λ, t) depends on λ in such wise that

$$\int_0^T \left[\frac{\partial A(\lambda, t)}{\partial \lambda}\right]^2 dt = \text{const.}$$

In this case, the error is completely determined by the function A(λ, t), and does not depend on P(λ). If, however, there exists a dependency

$$\int_0^T \left[\frac{\partial A(\lambda, t)^2}{\partial \lambda}\right]^2 \partial t = f(\lambda),$$

then, according to (10), the resulting error will depend on the parameter's probability distribution [the function P(λ)]. There now arises the problem of finding the optimal function A(λ, t) for a given P(λ).

The dependence $\int_0^T \left[\frac{\partial A(\lambda, t)}{\partial \lambda}\right]^2 dt = f(\lambda)$ can be obtained either by the use of a special "nonlinear" modulation or by a nonlinear parameter transformation prior to modulation.

We shall consider the second method in more detail.

In the transmitting device, before it is applied to the modulator, let the parameter λ be transformed to some auxiliary parameter y = y(λ). For transmitting the parameter y along the channel, such a form of modulation is used for which ∂A(y, t)/∂y does not depend on y, and the error in transmitting parameter y equals δ_0.

Then, by taking into account that ∂A(y, t)/∂λ = [∂A(y, t)/∂y] [∂y/∂λ], we find that the mean square error, averaged over all λ, equals

$$\bar{\delta}^2 = \frac{\sigma^2}{8\int_0^T \left[\frac{\partial A(y, t)}{\partial y}\right]^2 dt} \int_{-1}^{1} \frac{P(\lambda)}{y'^2} d\lambda = \delta_0^2 \int_{-1}^{1} \frac{P(\lambda)\, d\lambda}{y'^2},$$

where

$$y' = \frac{\partial y}{\partial \lambda}.$$

Thus, the gain due to the use of the nonlinear transformation equals

$$B_N = \frac{\delta_0^2}{\bar{\delta}^2} = \frac{1}{\int_{-1}^{1} \frac{P(\lambda)}{y'^2} d\lambda}. \tag{11}$$

The influence of the nonlinear transformation can also be explained without the ideal receiver. Figure 1 shows the nonlinear function y = y(λ) and the probability density P(λ). Let a quite small error Δy, constant over the entire range of values of parameter y, arise when the parameter y is transmitted.

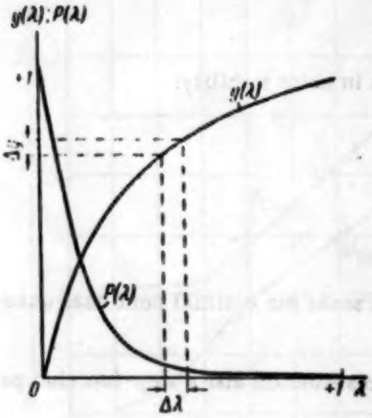


Fig. 1.

For small Δ, the errors Δy and Δλ are related by the relationship

$$\Delta y = \Delta\lambda y'.$$

The mean square error of reproduction of parameter λ equals

$$\overline{\Delta\lambda}^2 = \int_{-1}^{1} \frac{(\Delta y)^2 P(\lambda)}{y'^2} d\lambda.$$

Since $\Delta y = \delta_0$ and does not depend on λ, we then obtain the same result as in (11) for the ratio of the errors.

Obtaining the greatest gain in noise stability by taking statistics into account thus leads to the search for the optimal function $y = f(\lambda)$ which would minimize the integral

$$J = \int_{-1}^{1} \frac{P(\lambda)}{y'^2(\lambda)} d\lambda.$$

The problem thus posed is equivalent to the variational problem of finding the minimum of the integral

$$I(y) = \int_{x_1}^{x_2} F(\lambda,\ y,\ y')\, d\lambda.$$

Here, F is a continuous function, differentiable with respect to y' for all values of y' and differentiable with respect to the arguments λ and y in some region of the 0λy plane, and $y = y(\lambda)$ is a function which is continuously differentiable on the interval $\lambda_1 < \lambda < \lambda_2$. In our case, the function $F(\lambda,\ y, y') = P(\lambda)/y'^2(\lambda)$ satisfies all the conditions enumerated above.

The solution of such a variational problem leads to the differential equation

$$F_y - \frac{d}{\partial\lambda} F_{y'} = 0, \qquad (12)$$

where

$$F_y = \frac{\partial F}{\partial y}, \quad F_{y'} = \frac{\partial F}{\partial y'}.$$

In the case considered, $F_y = 0$, $F_{y'} = -2P(\lambda)(y')^{-3}$, and differential equation (12) has the form:

$$\frac{\partial P(\lambda)}{\partial\lambda}(y')^{-3} - 3P(\lambda)(y')^{-4} y'' = 0.$$

By assuming that $y' \neq 0$, which always holds in our practical problem, and by integrating the separated variables, we get

$$y' = c_0 \sqrt[3]{P(\lambda)}, \quad y = c_0 \int_{-1}^{\lambda} \sqrt[3]{P(\lambda)}\, d\lambda + c_1. \qquad (13)$$

The coefficients c_0 and c_1 are found from the conditions that y = ±1 for λ = ±1:

$$c_0 = \frac{2}{\int\limits_{-1}^{1} \sqrt[3]{P(\lambda)}\, d\lambda}, \quad c_1 = -1. \tag{14}$$

By substituting (13) and (14) in (11), we find the expression for the gain in noise stability:

$$B = \frac{2}{\left[\int\limits_{-1}^{1} P^{1/3}(\lambda)\, d\lambda\right]^{3/2}}.$$

It should be mentioned that an analogous solution is obtained when one seeks the optimal nonlinear quantizing of the signal [2].

We now consider two particular cases: a truncated normal probability distribution and a step-function parameter distribution.

The Truncated Normal Probability Distribution

Let P(x) satisfy equation (5), where q is defined by the condition

$$q \int\limits_{-1}^{1} e^{-\frac{\lambda^2}{2\lambda_0^2}}\, d\lambda = 1.$$

Then,

$$y' = c_0 \sqrt[3]{q}\, e^{-\frac{\lambda^2}{6\lambda_0^2}}.$$

The derivative of the optimal nonlinear function has the form of a normal law. The nonlinear transformation function sought will be the integral of the probability

$$y = c_0 \sqrt[3]{q} \int\limits_{-1}^{\lambda} e^{-\frac{\lambda^2}{6\lambda_0^2}}\, d\lambda - 1.$$

The gain in noise stability equals

$$B_N = \frac{0.175}{\lambda_0} \frac{\Phi^{1/2}(1/\lambda_0)}{\Phi^{3/2}(1/\lambda_0)},$$

where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{z} e^{-\frac{u^2}{2}}\, du.$$

Figure 2 gives the gain as a function of $1/\lambda_0$. For

$$\lambda_0 \ll 1 \quad B_N = 0.35\, \frac{1}{\lambda_0}.$$

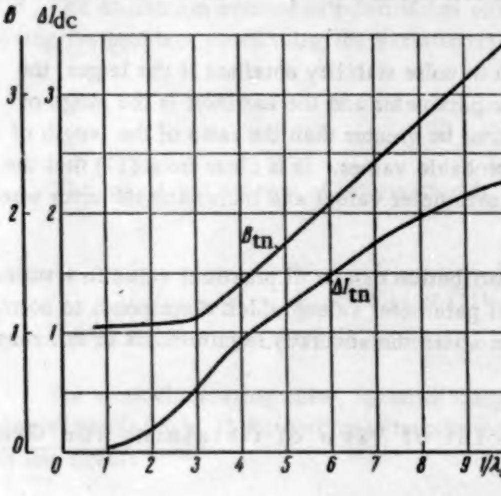


Fig. 2.

It is clear from the figure that an essential gain in noise stability (greater than a factor of 2) is only obtained for $1/\lambda_0 > 6$, i.e., in the case when the parameter dispersion is many times smaller than the entire range to be transmitted.

The Step-Function Probability Distribution

The step-function probability distribution (Fig. 3) may be written in the form

$$P(\lambda)=\begin{cases} R & \text{for} \quad |\lambda| \leqslant \frac{1}{\alpha}, \\ \frac{R}{\beta} & \text{for} \quad 1 \geqslant |\lambda| > \frac{1}{\alpha}. \end{cases} \tag{15}$$

If the parameters α and β are given, then the coefficient R can be determined from the condition

$$R\frac{2}{\alpha}+\frac{R}{\beta}\left(2-\frac{2}{\alpha}\right)=1,$$

from whence

$$R=\frac{\alpha\beta}{2(\alpha+\beta-1)}.$$

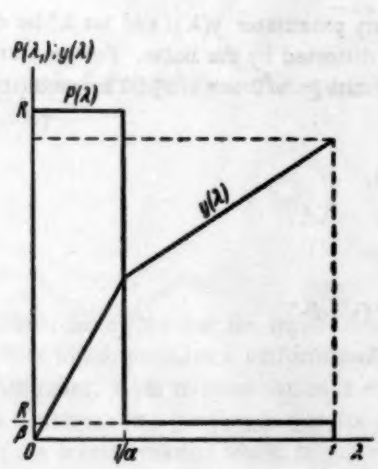


Fig. 3.

In accordance with (13) and (14),

$$y_I'=\frac{\alpha\sqrt[3]{\beta}}{\sqrt[3]{\beta}+\alpha-1} \quad \text{for} \quad |\lambda| < \frac{1}{\alpha},$$

$$y_{II}'=\frac{\alpha}{\sqrt[3]{\beta}+\alpha-1} \quad \text{for} \quad 1 > |\lambda| > \frac{1}{\alpha}.$$

The optimal nonlinear transformation for a step-function probability distribution is thus a piecewise linear transformation with breaks at the points $\lambda = \pm 1/\alpha$, where the slope of the line $y = f(\lambda)$ in the central portion is $\sqrt[3]{\beta}$ times larger than the slopes of the extreme portions.

The gain resulting from such a transformation equals

$$B_{st}=\alpha\frac{(\beta+\alpha-1)^{1/2}}{\left(\sqrt[3]{\beta}+\alpha-1\right)^{3/2}}. \tag{16}$$

For $\alpha \gg 1$ and $\beta \gg \alpha^3$, $B_{st} = \alpha$, i.e., the gain equals the ratio of the length of the entire range of the parameter to be transmitted to the length of the range of the probable parameter values.

It is easily shown that the ratio of the mean square error when transmitting parameter values in the range $|\lambda| > 1/\alpha$ to the error in the range $|\lambda| < 1/\alpha$ equals

$$\frac{\Delta_{II}}{\Delta_I}=\frac{y_I'}{y_{II}'}=\sqrt[3]{\beta}. \tag{17}$$

Example. Let $\alpha = 5$ and $\beta = 1000$. Then, the ratio of the errors equals $\Delta_{II}/\Delta_I = \sqrt[3]{1000} = 10$ and the gain in noise stability equals $B_{st} = 3.5$.

Thus, with a step-function probability distribution, the gain in noise stability obtained is the larger, the larger is the ratio of the probability densities in both ranges of the parameter and the narrower is the range of probable parameter values. It follows from (16) that the gain cannot be greater than the ratio of the length of the entire parameter range to the length of the range of its most probable values. It is clear from (17) that the gain is achieved by decreasing the error in the range of probable parameter values and increasing the error when low-probability parameter values are transmitted.

The case considered above of a step-function probability distribution can be of practical value in systems of telecommunications of parameters if there is a certain range of parameter values which corresponds to normal operation of the object and there is a range of abnormal operation where the accuracy requirements in this range are quite low.

A Nonlinear Transformation Optimal from the Point of View of Obtaining the Greatest Quantity of Information

We previously considered the question of a nonlinear transformation optimal from the point of view of minimizing the mean square error. We now consider this same problem from the information-theoretical point of view.

In a channel with noise, let there be transmitted some auxiliary parameter $y(\lambda)$, and let λ' be obtained by the inverse nonlinear transformation of the received signal $y_{rec}(\lambda)$, distorted by the noise. For greater clarity of the results, we will assume that the parameter λ varies within the limits [$-a/2$ to $+a/2$]. The quantity of information in λ' about λ equals

$$J = H(\lambda') - H_\lambda(\lambda').$$

Hence,

$$H(\lambda') = -\int_{-a/2}^{a/2} P(\lambda') \log P(\lambda')\, d\lambda'$$

and

$$H_\lambda(\lambda') = -\int_{-a/2}^{a/2} P(\lambda)\, d\lambda \int_{-a/2}^{a/2} P_\lambda(\lambda') \log P_\lambda(\lambda')\, d\lambda'. \tag{18}$$

For weak noise, we can assume that the error in the reproduced parameter is distributed normally. Due to the nonlinear transformation, the dispersion of this error is not constant over the range of values of λ:

$$\delta(\lambda) = \frac{\delta_0}{y'(\lambda)}.$$

With this,

$$P_\lambda(\lambda') = \frac{1}{\sqrt{2\pi}\,\delta(\lambda)} e^{-\frac{(\lambda-\lambda')^2}{2\delta^2(\lambda)}}. \tag{19}$$

By substituting (19) in (18) and by taking into account that the limits of the second integral in (18) can, without significant error, be replaced by infinity, we obtain

$$H_\lambda(\lambda') = \log \sqrt{2\pi e} - \int_{-a/2}^{a/2} P(\lambda) \log y'(\lambda)\, d\lambda.$$

The maximum amount of information will correspond to the minimum value of the quantity $H_\lambda(\lambda')$. By solving the problem posed using the variational method described above, we find that the minimum occurs for

$$y'(\lambda) = aP(\lambda). \tag{20}$$

The maximum amount of information is

$$J_{\max} = -\int_{-a/2}^{a/2} P(\lambda') \log P(\lambda')\, d\lambda' + \int_{-a/2}^{a/2} P(\lambda) \log P(\lambda)\, d\lambda + \log \sqrt{2\pi e}\, \frac{a}{\delta_0}.$$

For weak fluctuating noise, ignoring marginal values, we can assume that $P(\lambda) \approx P(\lambda')$. Then, $J_{\max} = \log \sqrt{2\pi e}\,(a/\delta_0)$. If the nonlinear transformation is not present, i.e., $y'(\lambda) = 1$, then the quantity of information equals

$$J = -\int_{-a/2}^{a/2} P(\lambda) \log P(\lambda)\, d\lambda - \log \sqrt{2\pi e}\, \delta_0.$$

The optimal nonlinear transformation thus allows one to transmit a quantity of information which is greater by the quantity

$$\Delta J = \log a + \int_{-a/2}^{a/2} P(\lambda) \log P(\lambda)\, d\lambda.$$

It follows from (20) that the transformation which is optimal from the point of view of quantity of information is that which provides a uniform probability distribution of the auxiliary parameter transmitted along the channel with noise. This transformation is different than in the case when the criterion was the mean square error, where it is nonlinear to a greater degree for the same law, $P(\lambda)$. This corresponds physically to the fact that, in the case of the transformation which is optimal from the information criterion, the rarer values of the transmitted parameter are transmitted with a larger error than in the case when the criterion was the mean square error.

Indeed, with the information criterion, the ratio of the errors for different λ is expressed by the formula

$$\left[\frac{\Delta(\lambda_1)}{\Delta(\lambda_2)}\right]_{\inf} = \frac{P(\lambda_2)}{P(\lambda_1)},$$

and for the mean square criterion

$$\left[\frac{\Delta(\lambda_1)}{\Delta(\lambda_2)}\right]_{\substack{\text{mn.}\\ \text{sq}}} = \frac{\sqrt[3]{P(\lambda_2)}}{\sqrt[3]{P(\lambda_1)}}.$$

The possible increase in the quantity of information for the truncated normal distribution [cf. (5)] equals

$$\Delta J_{\text{tn}} = \log \frac{1/\lambda_0}{\sqrt{2\pi}\,\Phi(1/\lambda_0)} - \left[1 + \frac{1/\lambda_0}{\sqrt{2\pi}\,\Phi(1/\lambda_0)} e^{-\frac{1}{2}\left(\frac{1}{\lambda_0}\right)^2}\right] \log \sqrt{e}.$$

The graph of this function is shown in Fig. 2.

For $1/\lambda_0 > 3$,

$$\frac{1}{\lambda_0} > 3 \qquad \Delta J_{\text{tn}} \approx \log 0.49 \left(\frac{1}{\lambda_0}\right).$$

In the case of the step-function distribution of probabilities (cf. (15)], the quantity of information can be increased via a nonlinear transformation by the amount

$$\Delta J_{\text{st}} = -\log \frac{\alpha + \beta - 1}{\alpha} + \frac{\beta}{\alpha + \beta - 1} \log \beta.$$

For $\beta \gg \alpha$ and $\alpha \gg 1$, we will have $\Delta J_{st} = \log \alpha$.

It is obvious that, for the nonlinear transformation which is optimal from the information criterion point of view, the mean square error will be greater than for the transformation which is optimal from the point of view of minimum mean square error. As computations show, this increase can be very significant. For example, in the case of a step-function distribution with $\beta = 1000$ and $\alpha = 10$, the quantity of information can be increased by $\log_2 10$. With this, the mean square error is increased by a factor of 100. This is due to the fact that the low-probability parameter values will be transmitted, with such a transformation, with an error 1000 times greater than for the most probable value. In practice, the choice of one criterion or another depends on the requirements imposed on the system for transmitting data. In certain cases both criteria considered here might be invalid, since the transmission of low-probability values with low reliability might be inadmissible in some operating conditions.

SUMMARY

1. If the relative dispersion of the parameter is much larger than the required accuracy, the use of the known first distribution law in constructing an ideal Kotel'nikov receiver provides little gain in noise stability. An essential disadvantage of the ideal receiver which takes the first probability law into account is the dependence of its mode of operation on the ratio of the noise spectral density to the parameter dispersion.

2. For any parameter probability distribution there may be found an optimal nonlinear transformation which minimizes the mean square error when the transmission is along a channel with weak fluctuating noise. In the cases where there is essential nonuniformity, the nonlinear transformations permit significant decreases in the mean square error.

3. To transmit the maximum quantity of information along a channel with weak fluctuating noise, the initial parameter probability distribution must be transformed to a uniform one. This means that the less frequent values will be transmitted with larger errors than the frequent ones. With this, the mean square error increases.

Received December 25, 1958

LITERATURE CITED

[1] V. A. Kotel'nikov, Theory of Potential Noise Stability [In Russian], Gosénergoizdat (1956).

[2] B. Smith, Instantaneous Compounding of Quantized Signals, The Bell System Technical Journal, No. 3, May, 1957.

ANALYSIS OF ROYER MULTIVIBRATOR OPERATION IN INDUSTRIAL TELEMETRY DEVICES

A. M. Pshenichnikov

(Moscow)

The conditions are investigated under which multivibrator oscillations are generated, the frequency of these oscillations being determined by the magnetic polarity reversal time of the core. The functional dependencies of the circuit's output frequency and voltage and its power requirements on the input voltage, the circuit parameters and the core material and transistor characteristics are determined.

A multivibrator circuit proposed by Royer [1] whose frequency is determined by the core's magnetic polarity reversal time is shown in Fig. 1. This circuit has been widely used in telemetering and analog computer devices as well as in power supplies for transforming dc to ac. Papers have recently appeared treating the use of Royer multivibrators in power supplies [2, 3].

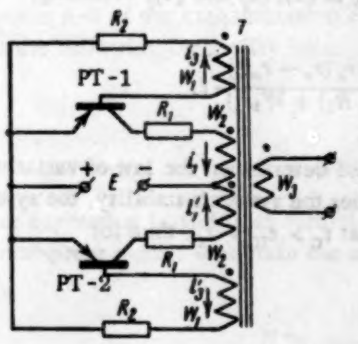


Fig. 1. A Royer multivibrator circuit.

In the present paper we investigate this circuit from the point of view of its use in telemetry devices. We consider the conditions under which oscillations are generated and we determine the circuit's output frequency and voltage and its power requirements as functions of the circuit parameters, the transistor characteristics and the core material characteristics.

For our analysis we approximate the hysteresis loop by straight-line seqments (Fig. 2). The magnetic permeability equals μ_0 on segments 1-2 and 3-4 and, on segments 2-3 and 4-1, is assumed initially to be $\mu = \infty$. In the sequel we shall determine the effect of a finite value of μ on multivibrator operation.

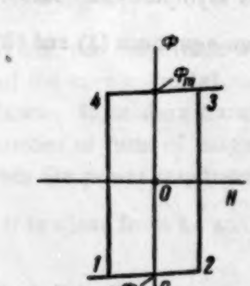


Fig. 2. The core's magnetization curve.

When a dc voltage is applied to the circuit's input, the collector current of one of the transistors (PT-1, say), due to the nonidentity of the transistor characteristics, will be greater than that in the other (PT-2). Thanks to this, the voltage induced in winding W_1 of transformer T will be so directed that minus is applied to the base of PT-1 and plus to the base of PT-2. This leads to a further increase in the collector current of PT-1 and a decrease in collector current in PT-2. As a result, PT-1 conducts completely and PT-2 is cut off. With this, the transformer core initially changes its magnetic polarity along arm a-2, and then along arm 2-3 (Fig. 2). The emf induced in winding W_1 falls sharply when point 3 on the magnetization curve is reached. The current in transistor PT-1 then decreases and, in winding W_1, an emf will be induced to the opposite sign. As the result, PT-1 is cut off and PT-2 conducts. Reversal of core magnetic polarity occurs along arms 3-4 and 4-1. Thereafter the circuit operates analogously. The multivibrator frequency is determined by the time taken to reverse magnetic polarity along arms 2-3 and 4-1, since $\mu >> \mu_0$.

I. Conditions for Oscillation Generation

Until oscillations are generated, the core's magnetic state is defined by point a on the hysteresis loop (Cf., Fig. 2). In order to simplify the analysis of the conditions for circuit excitation, we shall assume that, with fluctuating oscillations, current i_1 (Cf., Fig. 1) increases while current i_1' remains equal to zero. Thus, the problem reduces to the determination of the conditions for self-excitation of an amplifier with positive feedback, the equivalent circuit of which is shown in Fig. 3. In this circuit, r_e, r_b and r_c denote, respectively, the equivalent impedances of the emitter junction, the base and the collector and r_m denotes the equivalent junction impedance in the circuit with grounded base. The quantity $\alpha \approx r_m/r_c$ [4]. The equations for the circuit of Fig. 3 can be written as follows:

$$0 = i_b(r_e + r_b + R_2) + i_c r_e - W_2 \frac{d\Phi}{dt} 10^{-8}, \tag{1}$$

$$0 = i_b(r_e - r_m) + i_c(r_e + r_c - r_m + R_1) + W_2 \frac{d\Phi}{dt} 10^{-8}, \tag{2}$$

where i_b and i_c are, respectively, the base and collector current increments.

To determine the law of variation of the increment of current i_b we determined the quantity $d\Phi/dt$. When the circuit is switched in, the flux changes from point a along arm a-2 of the magnetization curve (Fig. 2) and its magnitude can be expressed in terms of the current increment in the following way:

$$\Phi = -\Phi_m + \frac{0{,}4\pi\mu_0 S}{l}(W_2 k - W_1) i_b, \tag{3}$$

where l is the mean length of the core's magnetic lines and $k = i_c/i_b$ is the gain. By differentiating equation (3) we find that

$$\frac{d\Phi}{dt} = \frac{0{,}4\pi\mu_0 S}{l}(W_2 k - W_1)\frac{di_b}{dt}. \tag{4}$$

After transformation, we solve the system of differential equations consisting of (1), (2) and (4), obtaining

$$i_b = i_{b0} \exp \frac{l\,[(r_b + r_e + R_2)(r_e + r_c - r_m + R_1) - r_e(r_e - r_m)]}{0{,}4\pi\mu_0 S\,(W_2 k - W_1)\,[W_1(r_e + r_c - r_m + R_1) + W_2 r_e]}\,t, \tag{5}$$

where i_{b0} is the initial deviation of the base current. The expression obtained determined the law of variation of small deviations of current from its value in the equilibrium state, and defines the system's stability, the system being unstable when the exponent is positive. Since it is always the case that $r_c > r_m > r_e$, then for

$$k > \frac{W_1}{W_2} \tag{6}$$

the size of the deviation grows without bound with increasing t and the system is unstable. For $k < W_1/W_2$ the size of the deviation tends to be zero with increasing t, and the equilibrium state is asymptotically stable [5].

The condition for system instability is found after the value of k obtained from equations (1) and (2), is substituted in expression (6)

$$-\frac{\frac{W_2}{W_1}(r_b + r_e + R_2) + r_e - \alpha r_c}{r_e + \frac{W_1}{W_2}[r_e + R_1 + r_c(1-\alpha)]} > 1. \tag{7}$$

It is clear from this inequality that to facilitate self-excitation, it is necessary to choose triodes with maximal α and to decrease impedances R_1 and R_2. To choose an admissible ratio, W_2/W_1, we solve the above

inequality for W_2/W_1. Ordinarily, W_1/W_2 is not chosen smaller than 0.5. Since $r_e \ll r_c(1-\alpha)$, then

$$r_e < \frac{W_1}{W_2}[R_1 + r_c(1-\alpha)].$$

After simplification of expression (7) and some transformation, we get

$$\left(\frac{W_2}{W_1}\right)^2 - \frac{\alpha r_c}{r_b + r_e + R_2}\,\frac{W_2}{W_1} + \frac{R_1 + r_c(1-\alpha)}{r_b + r_e + R_2} < 0. \tag{8}$$

By solving the quadratic inequality thus obtained, we find the permissible values of W_2/W_1

$$\frac{\alpha r_c - \sqrt{(\alpha r_c)^2 - 4(r_b + r_e + R_2)[R_1 + r_c(1-\alpha)]}}{2(r_b + r_e + R_2)} < \frac{W_2}{W_1} <$$
$$< \frac{\alpha r_c + \sqrt{(\alpha r_c)^2 - 4(r_b + r_e + R_2)[R_1 + r_c(1-\alpha)]}}{2(r_b + r_e + R_2)}. \tag{9}$$

For a triode with $\alpha = 0.9$, $R_C = 3\cdot 10^5$ ohm, $R_b = 200$ ohm, $r_e = 10$ ohm for $R_1 = 500$ ohm and $R_2 = 5000$ ohm, we have $0.03 < W_2 \; W_1 < 53$. For a triode with $\alpha = 0.98$, $r_c = 10^6$ ohm, $r_b = 400$ ohm, $r_e = 10$ ohm for $R_1 = 500$ ohm and $R_2 = 5000$ ohm we have $0 < W_2/W_1 < 172.5$.

Thus, as the transistor characteristics are improved, the range of admissible values of W_2/W_1 increases. However, there is no practical necessity to broaden the range of the ratio obtained from the first transistor.

Expression (7) is a necessary, but not sufficient, condition for the generation of oscillations by a Royer multivibrator. This is due to the fact that, although a small deviation of base current (5) will be increased if inequality (9) holds, it is not known whether or not, for given circuit parameters and source voltage E, the collector current reaches a value equal to the core's coercive force, i.e., whether or not equation (4) will be valid on the entire segment a-2 of the magnetization curve (Fig. 2). For equation (4) to be satisfied on segment a-2 it is necessary that the following inequality hold:

$$H_c < \frac{0.4\pi(i_c W_2 - i_b W_1)}{l}. \tag{10}$$

After expressing i_c and i_b in formula (10) in terms of the circuit parameters (Cf., Figs. 1 and 2), and after neglecting second-order terms, we obtain the second necessary condition for multivibrator operation

$$E \geqslant \frac{H_c l \alpha r_e r_c}{0.4\pi W_2[m r_c \alpha - m^2(R_1 + r_c - r_c\alpha) - r_b - r_e - R_2]}, \tag{11}$$

where $m = W_1/W_2$.

Thus, the magnitude of the minimum voltage necessary for circuit operation is proportional to the coercive force of the core material, to the mean length of the core's lines (mean diameter) and to the emitter junction impedance. In modern transistors the latter quantity varies comparatively insignificantly (from 20 to 40 ohms). The number of turns of magnetizing winding W_2 is given by the condition that the required frequencies be obtained and from the power requirements of the circuit (Cf., below).

It is clear from an analysis of formula (11) that the quantity E has a minimum for

$$m_k = \frac{\alpha r_c}{2(r_c - r_c\alpha + R_1)}. \tag{12}$$

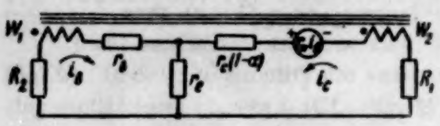


Fig. 3. Equivalent circuit for one of the amplifiers with positive feedback of the Royer multivibrator.

For the same coefficient m (for $m \le m_k$) the size of E decreases the increasing α, but this relation holds, basically, for values of m close to m_k. Variations in the base and collector impedances has little effect on the size of E.

II. Designing the Multivibrator Circuit

The problem of designing a complete multivibrator circuit (Fig. 4) is the determination of the dependence of the output frequency and voltage and of the multivibrator power requirements on the input voltage, on the circuit parameters and on the characteristics of the core material.

For simplifying the computations we shall assume that the multivibrator's output frequency is determined only by the core's magnetic polarity reversal, and does not depend on the transient response connected with the switching of the transistors. The latter assumption is valid since, for the frequencies considered here, the processes related to the switching of the transistors occur much more rapidly than the core's magnetic polarity reversal. In the multivibrator circuit of Fig. 4, in contradistinction to the basic circuit (Fig. 1), there have been added regulating resistors R_3 and R_4, made by tapping off a portion of the magnetizing winding with number of turns equal to W_2', for the purpose of regulating the output frequency, and by the inclusion of an additional supply voltage E_1. With our assumptions taken into account, we consider the circuit of Fig. 5, equivalent to half the multivibrator. In this circuit, the transistor is replaced by key K. The relationship between the circuit parameters and the time of flux growth in the core can be obtained from the following system of equations, set up for the circuit of Fig. 5 in accordance with Kirchoff's laws:

$$E = iR_1 + i_2R_4 + W_2\frac{d\Phi}{dt}10^{-8}, \tag{13}$$

$$i_1R_3 + W_2'\frac{d\Phi}{dt}10^{-8} = i_2R_4, \tag{14}$$

$$i = i_1 + i_2, \tag{15}$$

$$W_3\frac{d\Phi}{dt}10^{-8} + i_LR_L = 0, \tag{16}$$

$$E_1 = i_3R_2 - W_1\frac{d\Phi}{dt}10^{-8}, \tag{17}$$

$$H_f + H_1 + H_2 + H_2' + H_3 = H_c. \tag{18}$$

Here, H_f is the field strength induced by the eddy currents, H_1 is the field strength induced by current i_3 flowing in winding W_1; H_2 is the field strength induced by current i flowing in winding W_2; H_2' is the field strength induced by current i_1 flowing in winding W_2'; H_3 is the field strength induced by current i_L flowing in winding W_3. Equation (18) holds as a consequence of the equality of the total field strength induced by currents i, i_1, i_3, i_L and the eddy currents during the entire transient response and, on the other hand, the coercive force (with reversal magnetic polarity along portion 2-3 of the magnetization curve of Fig. 2).

By solving the system of equations (13-18) we obtain, after some transformation, the differential equation

$$\begin{aligned}\frac{0.4\pi\cdot10^{-8}}{l}\left\{\frac{W_2^2(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]}+\frac{W_1^2}{R_2}+\frac{W_3^2}{R_L}\right\}d\Phi+\\+0.8\cdot10^{-8}\pi B_m\gamma z\,dz=\left\{\frac{0.4\pi W_2(1+a)(1+n+an)}{nlR_1[(1+a)^2+ac]}E-\frac{0.4\pi W_1E_1}{lR_2}-H_c\right\}dt,\end{aligned} \tag{19}$$

where $R = R_3 + R_4$, $a = R_3/R_4$, $n = W_2/W_2'$, $c = R/R_1$, γ is the specific conductivity of the core material, B_m is the saturation inductance, the quantities z and dz are shown on Fig. 6 [6]. As the flux changes from $-\Phi_m + H_c\mu_0S$ to $\Phi_m + H_c\mu_0S$, where S is the core's cross-sectional area, the time changes from 0 to T/2 and z changes from 0 to d/2, where d is the thickness of the Permalloy strip.

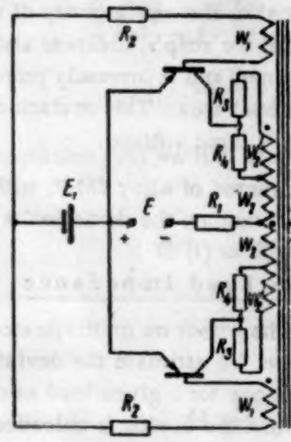


Fig. 4. Multivibrator with frequency tuning by changing of the ratio R_3/R_4.

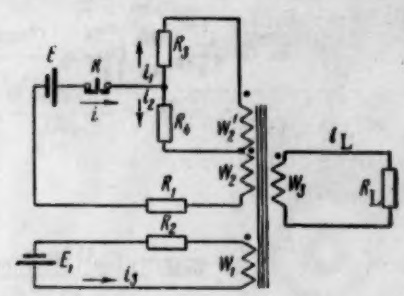


Fig. 5. Functional schematic for investigating multivibrator core magnetization polarity reversal.

To determine the multivibrator's output frequency, we integrate expression (19)

$$f = \frac{1}{T} = \frac{\dfrac{0{,}4\pi W_2(1+a)(1+n+an)}{nlR_1[(1+a)^2+ac]}E - \dfrac{0{,}4\pi W_1 E_1}{lR_2} - H_c}{\dfrac{1{,}6\pi\cdot 10^{-8}\Phi_m}{l}\left\{\dfrac{W_2^2(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]} + \dfrac{W_1^2}{R_2} + \dfrac{W_3^2}{R_L}\right\} + \pi\cdot 10^{-9}B_m\gamma d^2} \tag{20}$$

Formula (20) shows that, for a multivibrator constructed according to the circuit of Fig. 4 with the core material's coercive force and the eddy currents taken into account, and with the magnetic permeability equal to μ_0 on portions 1-2 and 3-4 of the magnetization curve and equal to $\mu = \infty$ on portions 2-3 and 3-4, the pulse frequency is proportional to the input voltage and inversely proportional to the saturation inductance and (to a first approximation) to the core's cross-sectional area ($\Phi_m = B_m S$).

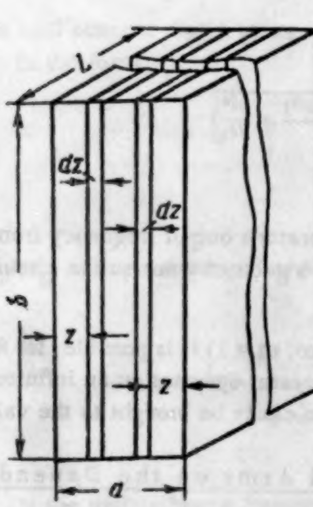


Fig. 6. For taking into account the effect of the eddy currents during magnetic polarity reversal.

1. Effect of the Eddy Currents on the Frequency

The effect of the Foucault currents on the frequency can be defined as the deviation of the output frequency f_1, determined without the eddy currents being taken into account, from the frequency f determined with the latter taken into account

$$\delta\% = \frac{f_1 - f}{f}100\%.$$

By substituting, instead of f_1 and f, the corresponding expressions obtained from formula (20), we obtain, after some transformation,

$$\delta\% = \frac{\gamma d^2 l}{16S\left\{\dfrac{W_2^2(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]} + \dfrac{W_1^2}{R_2} + \dfrac{W_3^2}{R_L}\right\}}100\%. \tag{21}$$

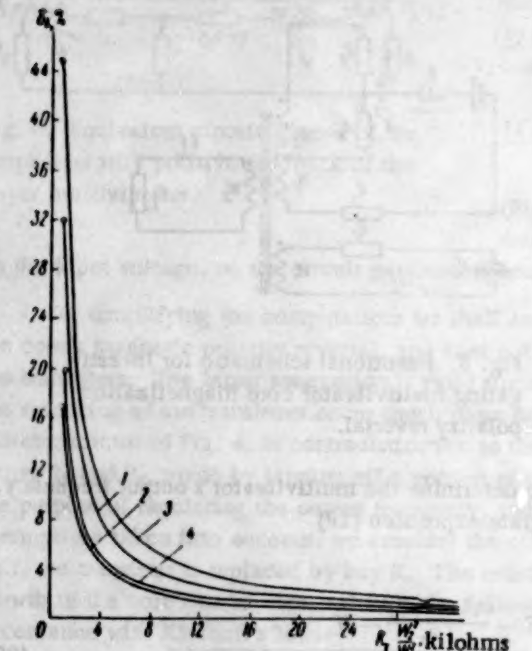


Fig. 7. Deviation of the frequency for a given reduced load impedance on the multivibrator's output from the unloaded frequency. 1) a = 0, m = 1; 2) a = 0, m = 5; 3) a = ∞, m = 1; 4) a = = ∞, m = 5, n = 5, c = 2, R = 1000 ohm, R = 500 ohm.

It is clear from expression (21) that the deviation of the frequency from the true value when one ignores the eddy currents is proportional to the conductivity of the core material, the square of the strip's thickness and the mean length of the core's lines, and is inversely proportional to the core's cross-sectional area. This deviation does not depend on the circuit's input voltage.

For cores constructed of alloy 65NP, with a strip thickness of 0.05 millimeters, the deviation is about 0.5%.

2. Effect of the Load Impedance

To determine the effect on multivibrator operation of the load impedance, we estimate the deviation of the device's output frequency for a given load as compared to the frequency, f_2, when the device is unloaded

$$\delta_1\% = \frac{f_2 - f_1}{f_1} 100\%.$$

If, instead of f_1 and f_2, we substitute the corresponding expressions from formula (20), and after some transformation, where

$$R_{\mathrm{L\,re}} = R_{\mathrm{L}}\left(\frac{W_2}{W_3}\right)^2,$$

we obtain

$$\delta_1\% = \frac{100\%}{R_{\mathrm{L\,re}}\left\{\frac{(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]} + \frac{m^2}{R_2}\right\}}. \qquad (22)$$

Thus, as the load impedance R_L varies, the deviation of the multivibrator's output frequency from the unloaded frequency varies along a hyperbola and depends neither on the core's geometry nor on the circuit's input voltage. Figure 7 shows δ_1% as a function of R_{Lre}.

It is clear from the curves of Fig. 7 that even in the worst case (a = ∞, m = 1) it is possible, for R_{Lre} = = 15,000 ohm, to assume, with sufficient design accuracy, that the multivibrator operates on an infinite load (free running). With an output amplifier the multivibrator's load impedance can easily be brought to the value cited.

3. Effect of the Slope of the Hysteresis Loop's Lateral Arms on the Dependence on Voltage of the Multivibrator's Output Frequency

Formula (20) was derived from the condition that the lateral arms of the hysteresis loop were vertical. This is defined by equation (18). If the slope of the lateral arms of the hysteresis loop (Cf., Fig. 8) is taken into account then, in the system of equations consisting of (13-18), equation (18) must be replaced by the following relationship:

$$H_f + H_1 + H_2 + H_2' + H_3 = \frac{\Phi + H_c \mu S}{\mu S}. \qquad (23)$$

If we neglect, for simplicity of the computations, the action of the eddy currents and the effect of the load resistance, which is possible if the conditions enumerated above hold, then we can solve the system of equations (13-17) and (23) to obtain

$$\frac{0.4\pi\cdot 10^{-8}}{l}\left\{\frac{W_2^2(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]}+\frac{W_1^2}{R_2}\right\}\frac{d\Phi}{dt}+\frac{1}{\mu S}\Phi=$$
$$=\frac{0.4\pi W_2(1+a)(1+n+an)}{nlR_1[(1+a)^2+ac]}E-\frac{0.4\pi W_1E_1}{lR_2}-H_c. \tag{24}$$

By solving equation (24) we find that

$$\Phi(t)=M\mu S-\left[M\mu S+\Phi_m+\mu_0\frac{\Phi_m-\mu SH_c}{\mu-\mu_0}\right]e^{-\frac{t}{M\mu Q}}, \tag{25}$$

where

$$M=\frac{0.4\pi W_2(1+a)(1+n+an)}{nlR_1[(1+a)^2+ac]}E-\frac{0.4\pi W_1E_1}{lR_2}-H_c,$$

$$Q=\frac{0.4\pi\cdot 10^{-8}}{l}\left\{\frac{W_2^2(1+a)(1+a+c+2cn+cn^2+acn^2)}{cn^2R_1[(1+a)^2+ac]}+\frac{W_1^2}{R_1}\right\}.$$

The time necessary to reverse the core's magnetic polarity from point 2 to point 3 on the magnetization curve (cf. Fig. 8) is found from the formula

$$\frac{T}{2}=-\mu SQ\ln\frac{M\mu S-\Phi_m-\mu_0\dfrac{\Phi_m-\mu SH_c}{\mu-\mu_0}}{M\mu S+\Phi_m+\mu_0\dfrac{\Phi_m-\mu SH_c}{\mu-\mu_0}}. \tag{26}$$

If we take into account that $\mu \gg \mu_0$, $M > H_c$ and $B_m = \Phi_m/S$, we can write the expression for the output frequency in the form

$$f=\frac{1}{2\mu SQ\left[\ln\left(1+\frac{B_m}{M\mu}\right)-\ln\left(1-\frac{B_m}{M\mu}\right)\right]}. \tag{27}$$

Since $B_m/M\mu < 1$ we obtain, after some transformations,

$$f=f_\infty\frac{1}{1+\frac{1}{3}\left(\frac{B_m}{M\mu}\right)^2+\frac{1}{5}\left(\frac{B_m}{M\mu}\right)^4+\cdots+\frac{1}{2p+1}\left(\frac{B_m}{M\mu}\right)^{2p}}, \tag{28}$$

where f_∞ is the multivibrator frequency for $\mu=\infty$, and $\frac{1}{2p+1}\left(\frac{B_m}{M\mu}\right)^{2p}$ is the general term of the series.

The deviation of the multivibrator's output frequency when the magnetic permeability is μ as compared to the frequency when $\mu=\infty$ is defined by the expression

$$\delta_2=\frac{f-f_\infty}{f_\infty}100\%=-\frac{\frac{1}{3}\left(\frac{B_m}{M\mu}\right)^2+\frac{1}{5}\left(\frac{B_m}{M\mu}\right)^4+\cdots+\frac{1}{2p+1}\left(\frac{B_m}{M\mu}\right)^{2p}}{1+\frac{1}{3}\left(\frac{B_m}{M\mu}\right)^2+\frac{1}{5}\left(\frac{B_m}{M\mu}\right)^4+\cdots+\frac{1}{2p+1}\left(\frac{B_m}{M\mu}\right)^{2p}}100\%. \tag{29}$$

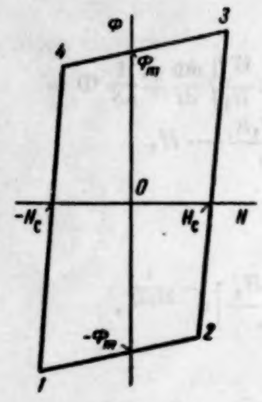


Fig. 8. Magnetization curve.

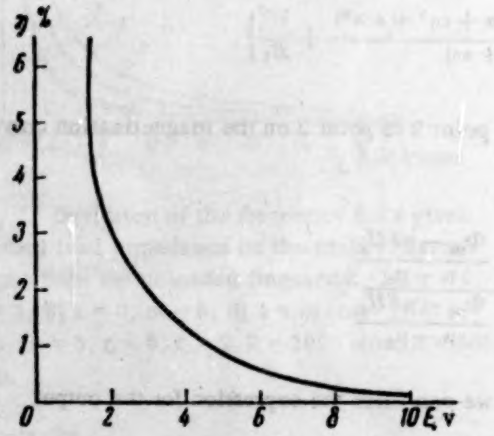


Fig. 10. Dependence on the applied voltage of the error in determining δ_2 from the simplified formula, for H_c = 0.2 oersted and μ = 50,000.

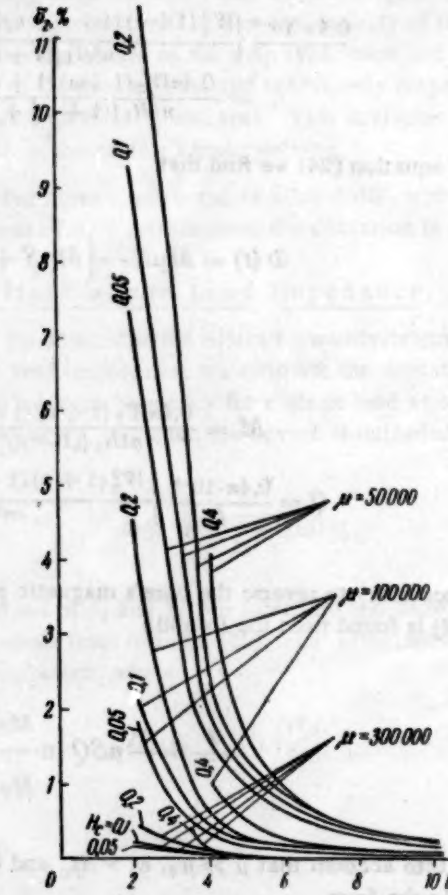


Fig. 9. Dependence on the magnitude of the input voltage of the deviation of the multivibrator's output frequency when the magnetic permeability is μ and the coercive force is H_c from the frequency for an ideal rectangular hysteresis loop.

Figure 9 shows the dependence of the deviation, δ_2%, on the applied voltage for various μ and H_c, computed from formula (29) for a multivibrator containing a core whose mean line length is l = 11.3 centimeter, B_m = 10,000 gauss, for W_1 = 500, W_2 = 900, R = 1000 ohm, $R_L = \infty$, R_2 = 10,000 ohm, c = 2, a = 0, n = 4.5 and E = 0.4 volt. With adequate design accuracy, and taking into account that $B_m M_\mu < 1$, we can determine the deviation from the simplified formula

$$\delta_2'\% = -\frac{B_m^2}{3(M\mu)^2}100\%. \tag{30}$$

Figure 10 shows the dependence of the error in determining δ_2% from formula (30) instead of formula (29) for p = 2

$$\eta\% = \frac{\delta_2' - \delta_2}{\delta_2}100\% \tag{31}$$

for H_c = 0.2 oersted and μ = 50,000.

As is obvious from the graphs, the maximum value of η does not exceed 6.5%. Thus, the size of δ_2, with sufficient accuracy, can be considered proportional to the square of the residual inductance and inversely proportional to the square of the magnetic permeability and the square of the applied voltage.

In calculating the error in a telemetry system, it is desirable to know the error reduced to a nominal scale. This error can be calculated from the formula

$$\delta_{2\,\mathrm{re}}\,\% = \frac{E}{E_N}\,\delta_2\% \tag{32}$$

where E_N is the nominal input voltage to the circuit.

It is clear from the analysis provided that for large residual inductances and for $\mu < 10^5$ there arises an essential nonlinearity in the relationship between the output frequency and the input voltage. When alloy 65NP is used, the error from this nonlinearity does not exceed 0.1%, while for alloy 50NP it is less than 1%. This figures are valid only for well-annealed cores.

4. Influence of the Impedance of the Magnetizing Circuit and of the Coefficient on a Circuit Operation

For stable circuit operation, it is important that the operation be independent of the variations in the magnetizing circuit impedance R_1, a contribution to whose magnitude is also made by the emitter-collector impedance of the conducting triode. As shown by analysis of circuit operation, the dependence of output frequency on the impedance R_1 will be at its greatest when a = ∞. However, even in this case, as shown by an analysis of formula (20), the output frequency will not depend on R_1 if the following three inequalities hold:

$$\begin{gathered}\frac{0.4\pi}{lR_1}E \geqslant \frac{0.4\pi m E_1}{lR_2} + \frac{H_c}{W_2},\\ \frac{1}{R_1} \gg \frac{1}{n^2R} \qquad \frac{1}{R_1} \gg \frac{m^2}{R_2}.\end{gathered} \tag{33}$$

It follows from an analysis of the inequalities in (33) that, for independence of circuit operation from variations in impedance R_1, one should decrease R_1 as much as possible, increase impedance R_2 and choose a core material with the minimum coercive force. The coefficient $m = W_1/W_2$ should be chosen with simultaneous consideration being given to the conditions of decreasing the minimum voltage necessary for generation of oscillations and the independence of circuit operation from variations in impedance R_1. These requirements contradict one another, and so a compromise solution must be reached.

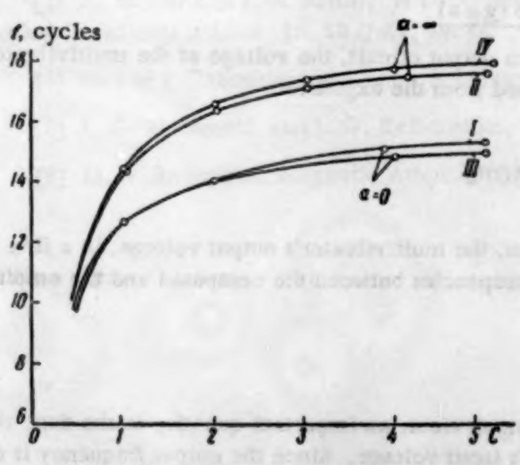


Fig. 11. Experimental and computed curves of the multivibrator frequency as a function of the coefficient c for a = 0 and a = ∞.

Figure 11 gives the experimentally obtained dependence of the frequency on the coefficient c for a multivibrator with a = 0 and with a = ∞ (curves I and II). The same figure gives the dependencies of the frequency on the coefficient c as determined from formula (20), namely, curves III and IV.

The discrepancies between the computed and the experimental data do not exceed 2 to 3%.

It is clear from the graphs of Fig. 11 that a variation of the coefficient a from 0 to ∞, i.e., a translation of the potentiometer slider from one extreme position to the other, can change the multivibrator frequency by ±8 to 10%. This is particularly important for mass production, when the characteristics of all the transmitting telemetry devices must be identical and not depend on the variations in core material characteristics.

5. Multivibrator Drain on the Voltage Source to be Measured

The average current drawn by the circuit is

$$I_{av}=\frac{2}{T}\int_0^{T/2} i\,dt, \tag{34}$$

where i is the value of current defined by equations (13-18)

$$i=\frac{En(1+a)-W_2(1+n+an)\frac{d\Phi}{dt}10^{-8}}{nR_1\left(1+a+\frac{ac}{1+a}\right)} \tag{35}$$

By substituting (35) in (34) we obtain, after integration,

$$I_{av}=\frac{E(1+a)}{R_1\left(1+a+\frac{ac}{1+a}\right)}-\frac{4W_2\Phi_m f(1+n+an)10^{-8}}{nR_1\left(1+a+\frac{ac}{1+a}\right)}. \tag{36}$$

In formula (36) the frequency value substituted for f should be computed from formula (20) with account taken of the load impedance and the Foucault currents. Figure 12 gives the values of average current drawn by the multivibrator as computed from formula (36) (curve 1). The discrepancies of the computed data from the experimental (curve 2) are about 5 to 7%.

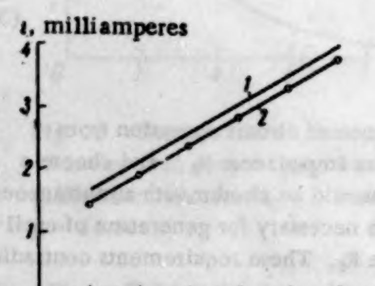


Fig. 12. Current drawn by the multivibrator as a function of applied voltage. 1) is the computed function and 2) is the experimentally determined function.

It is clear from an analysis of formulae (36) and (20) that the current drawn by the circuit is proportional to the input voltage and the coercive force of the core material. The basic method of reducing the drain of the multivibrator on the transducer amounts to decreasing the voltage E necessary for obtaining a given nominal frequency. However, the circuit must, with this, operate stably with a minimal frequency. Obviously, the minimum voltage at the circuit's input must satisfy condition (11). Starting from this condition, one must decrease the core's diameter and increase the number of turns, W_2.

6. The Multivibrator's Output Voltage as a Function of the Input Signal

With an open output circuit, the voltage at the multivibrator's output is determined from the expression

$$E_{out}=W_3\frac{d\Phi}{dt}\cdot 10^{-8}. \tag{37}$$

The magnitude of $d\Phi/dt$ can be found from equation (19). Thus, the multivibrator's output voltage, in a first approximation, depends linearly on its input voltage. The discrepancies between the computed and the empirical characteristics do not exceed 15%.

7. Speed of Response

When a multivibrator is used in telemetering or computing devices, an important quantity is the time taken to establish its output frequencies after a change in the circuit's input voltage. Since the output frequency is determined by the time taken to reverse the magnetic polarity of the multivibrator's transformer core, the transient response duration does not exceed half a period of the newly established frequency corresponding to the new voltage value.

SUMMARY

1. The multivibrator's pulse frequency is strictly proportional to the circuit's input voltage only in the case when the magnetic permeability on the hysteresis loop's lateral arms (segments 2-3 and 4-1 on Fig. 2) equals ∞. Otherwise, there is a nonlinearity error. This latter error is about 0.1% when well-annealed cores of alloy 65NP are used, and about 1% when cores of alloy 50NP are used.

2. Eddy currents decrease the multivibrator output frequency but do not distort the linear relationship between frequency and input voltage.

3. The variation in multivibrator frequency when the load resistance is lowered, as compared with the free-running frequency, is inversely proportional to the reduced load resistance and does not depend on the core geometry or on the circuit's input voltage.

4. With the proper choice of parameters, the operation of the circuit depends very little on variations in the impedance of the magnetizing circuit, a contribution to which is made by the emitter-collector impedance of the conducting transistor.

5. For tuning purposes, the multivibrator's output frequency can be varied by 8 to 10% by moving a potentiometer slider (by changing the coefficient $\underline{a}$ from 0 to ∞).

6. The current drawn by the multivibrator from the transducer is proportional to the input voltage and to the coercive force of the core material.

7. The time to establish the output frequency, when the signal at the multivibrator's input changes, does not exceed half a period of the frequency corresponding to the new voltage value.

8. The core material must possess minimal coercive force, a high value of maximum magnetic permeability and of saturation inductance and an adequate specific resistivity. Alloy 65NP fully satisfies these requirements.

Received February 25, 1959

LITERATURE CITED

[1] G. H. Royer, A Switching Transistor DC to AC Converter Having an Output Frequency Proportional to the DC Input Voltage, AIEE Transaction, vol. 74, pt. I, July, 1956.

[2] G. N. Brestovskii and K. Ya. Senatorov, "A voltage transformer made of power transistors," [in Russian], Radiotekhnika i Élektronika 9 (1957).

[3] A. Muller and J. de Saltre, "A converter from dc made of transistors mounted in push-pull," [in French] Annales de radioelectricité 13. 53 (July, 1958).

[4] Sh and R Transistors and Their Applications, Gozénergoizdat, (1957).

[5] L. G. Meerovich and L. G. Zelichenko, Pulse Technology [in Russian], Sovetskoe Radio (1956).

[6] M. A. Rozenblat, Magnetic Amplifiers [in Russian], Sovetskoe Radio (1956).

FREQUENCY ELEMENTS OF REMOTE CONTROL

I. N. Lisitskaya, V. N. Mikhailovskii, K. D. Nadtochii
and V. N. Okhotskaya

(L'vov)

Certain types of frequency relays are considered. Data are provided on relays and frequency generators developed at the Institute for Machinery (mashinovedeniya) and Remote Control of the AN USSR.

The majority of dispersed objects subject to remote control (irrigation systems, gas and oil fields, gas pipelines etc.) are generally interconnected by two-conductor telephone lines. When such objects are subjected to remote control, it is reasonable to make use of frequencies above and below the audio range, since this makes it possible to have parallel operation of the telephonic communications link and the remote control apparatus. It is more efficient (economically) to operate in the sub-audio frequency range ($f \leq 300$ cycles), where the signal damping in the conductors is relatively small, which permits the use of apparatus over tens and even hundreds of kilometers while still guaranteeing noise stability. One of the principal problems which arises with remote control (dispatcher control) of dispersed objects is the choice of method and means of selecting an object.

By using the very promising frequency-combination method of selection with time relaying [1] for a small number of working frequencies ($n \leq 10$), well spaced in the chosen range ($f \leq 300$ cycles), and by using $r \leq 3$ nonrepeating frequencies for each point, one can provide reliable selection of the necessary number of objects ($N \leq 720$) joined by parallel lines in which there is noise from such sources as magneto ringing of the telephone apparatus and lightening discharge. To implement this, it is necessary to have a highly stable characteristic matching of the transmitting generator and the frequency filter (relays). This matching must amount to this, that the possible bias, δ, in the transmitting generators frequency, f_{gk}, must be less than half the transmittance band of the filter, and the pass bands of filters with neighboring frequencies do not overlap.

Naturally, one tries to arrange matters so that the operating range of signal voltages and the external temperature are as large as possible, the filters' input impedance is as high as possible $R_{in} \geq 1$ kilohm), the dimensions and the power dissipation are as low as possible, and the output power is sufficient for reliable operation of the simplest and cheapest small-dimension telephonic relay, for example, type RKM ($P \gtrsim 0.1$ watt).

The conditions just posed may be satisfied quite adequately by relays and master generators with electromechanical filtering elements, and even by crystal-triode relays with RC-filters. The most promising of the electromechanical filters are the reed filters. Slide-wire filters are difficult to build and have very low efficiencies, and tuning-fork filters are complicated to prepare and their quality is unjustifiably high for the conditions considered, which require only a strict stabilization of the master generators' frequencies.

Today, even the best of the well-known electromechanical frequency relays with reed filters possess certain design complexities, relatively long operation and release times, and sizeable dimensions. In this paper we shall attempt to investigate the possibilities of creating frequency elements with improved parameters.

1. Contactless Frequency Elements Based on Reed Filters

An electromechanical filter, as is well-known, works on the principle of electromechanical resonance of a steel lamina oscillating in the field of a constant magnet under the action of the variable magnetic flux of an

excitation winding. The resonant frequency of the electromechanical filter, not being the frequency of the lamina's free oscillations, depends not only on the elasticity P_0 and mass $\underline{m}$ of the lamina, but also on the magnitude of the constant magnets' inductance in the air gap, B_0, and in the gaps δ_1 and δ_2 (in the lower and upper portions of the magnetic system) and on the ampere-turns, Iw, of the excitation winding.

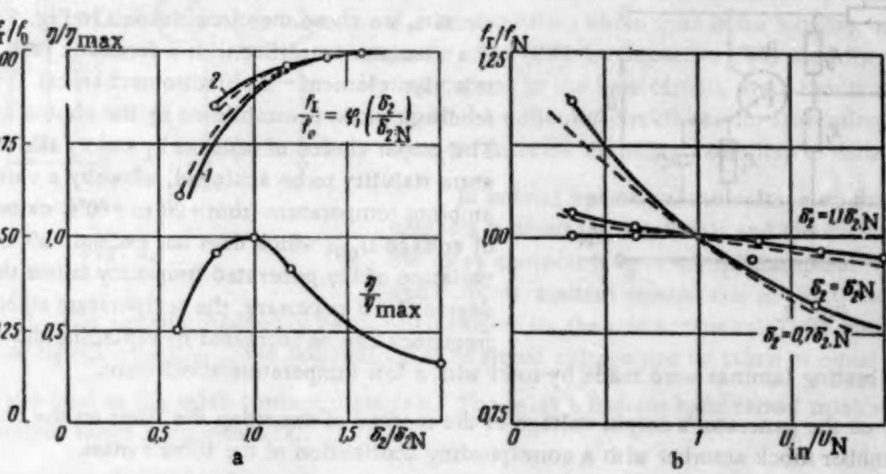


Fig. 1.

The expression for determining the resonant frequency of an electromechanical vibrator can be obtained from the equation [2.3]

$$\frac{d^2x}{dt^2} + \frac{k_d}{m}\frac{dx}{dt} + \left[\omega_0^2 - \frac{B_0^2}{\delta_1+\delta_2}K - \frac{B_0 Iw}{(\delta_1+\delta_2)^2}M\right]x = \frac{B_0 Iws}{\delta_1+\delta_2}\sin\omega t, \tag{1}$$

where $\underline{x}$ is the deviation of the lamina from its mean position, k_d is the damping coefficient, K and M are coefficients which depend on the parameters of the electromechanical vibrator , $\underline{s}$ is the cross-sectional area of the air gap and I is the current in the excitation winding.

It is easy to find the expression for the electromechanical vibrator's resonant frequency from equation (1):

$$f_r = \frac{1}{2\pi}\sqrt{\omega_0^2 - \frac{B_0^2}{\delta_1+\delta_2}K - \frac{B_0 Iw}{(\delta_1+\delta_2)^2}M}. \tag{2}$$

Expression (2) shows that increasing the induction B_0 and the current in the excitation winding, I, leads to a decrease of the resonant frequency f_r. When B_0 and I are decreased, the resonant frequency of the electromechanical vibrator increases, approximating to the value of the lamina's natural frequency of oscillation ($f_r \to f_0$ for $B_0 \to 0$ and $I \to 0$). Change of the air gap also leads to a variation in f_r. When the gap is increased the frequency f_r increases, and tends to f_0, but the dependence of the frequency on the magnitudes of B_0, I and δ is decreased. Figure 1 gives the analytically obtained functions $f_r/f_0 = \varphi(\delta_2/\delta_{2N})$ for U_{in} = const, and $f_r/f_N = \varphi_2(U_{in}/U_N)$ for δ_2 = const (where δ_{2N} is the size of the magnetic gap for which the efficiency η is maximal, and U_{in}/U_N is the ratio of the input signal's voltage to the voltage taken as nominal), which coincide well with the experimentally determined characteristics, shown by the dashed lines.

For carrying out experiments we prepared specimens of the electromagnetic filters for frequencies of 60-300 cycles, the laminas having a thickness of 0.35 mm, a width of 88 mm and lengths of 30 to 40 mm. With this, the size of the gap was $\delta_{2N} \approx 0.5$ mm. The filter dimensions were 25 × 35 × 60 mm.

The characteristics given in Fig. 1 show the decrease in the dependency of the electromechanical vibrator's frequency on the instability of the gaps and the input signal levels as the gaps δ_1 and δ_2 are increased. Curve 1 was obtained for $\delta_1 = 0$ and curve 2 for $\delta_1 = \delta_{2N}$. To increase the stability of the electromechanical filter's

resonant frequency (which is particularly essential for a frequency generator) it is necessary to guarantee that it operates on the section of the curve for which $\delta_2 / \delta_{2N} > 1$.

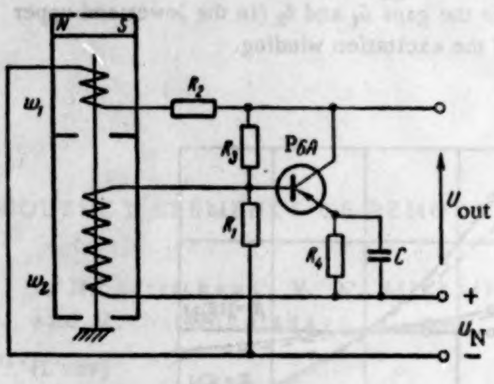


Fig. 2.

As a result of the investigation of several generator circuits, we chose the circuit shown in Fig. 2, consisting of a transistor amplifier with a feedback path through a selection element – an electromechanical vibrator. The feedback mode is established by the choice of R_2, R_3 and C. The proper choice of resistors R_1 and R_4 allows high temperature stability to be achieved, whereby a variation of the ambient temperature from +20 to +60°C causes a variation of voltage U_{out} which does not exceed ±5%, and the variation of the generated frequency is less than 0.01% per degree C. If necessary, the temperature stability of the frequency can be increased by replacing the ordinary spring steel,of which the vibrating laminas were made by steel with a low temperature coefficient.

The influence on the generator's output voltage of the method of mounting the filter on the chassis was avoided by using a rubber shock absorber with a corresponding distribution of the filter's mass.

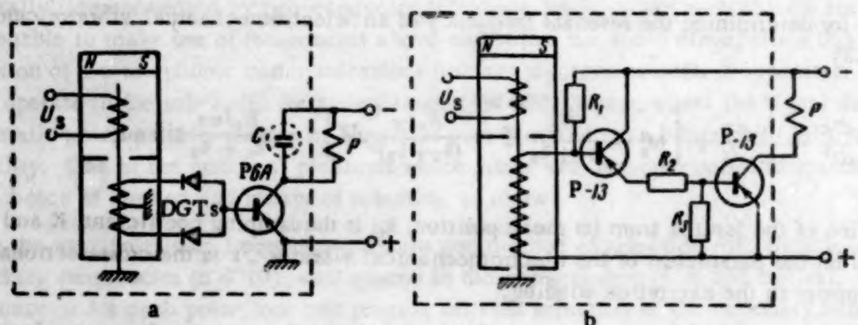


Fig. 3.

The best of the contactless electromechanical frequency relay circuits investigated turned out to be the ones shown in Fig. 3,a and b*. In the circuit of Fig. 3,b a cathode follower was introduced in order to match the impedances of the electromechanical filter's winding and the output amplifier.

The input power of the frequency relay necessary for reliable operation of relay RKM, connected at the output, (R_{wi} = 700 ohm, P_{out} = 0.15 to 0.2 watt) was 5 milliwatts. The relay's input impedance was $R_{in} \gtrsim 3000$ ohms. The nominal input voltage was $U_{N.s}$ = 5 to 7 volts, the relative working frequency band, $2\Delta f_{av}(100\%)/f_{n\ av}$, was 8 to 12% when $U_s = 1.3\ U_{N.s}$ and 3 to 5% when $U_s = 0.7\ U_{N.s}$.

When the ambient temperature varied from 0 to 40°C there was an insignificant (≤0.5%) narrowing of the working frequency band. There was no relay mistuning with variations in temperature.

The working range of the frequency relay's input signal voltage and, consequently, the reliability of the selection system operation under variation in communication channel parameters can be broadened by a factor of from 5 to 10 by connecting a signal amplitude limiter (for example, one composed of four crystal diodes) at the relay's input. In the frequency range ΔF = 45 to 315 cycles, we succeeded relatively easily in positioning nine or ten frequency channels. The physical construction of the filters was essentially simplified by the replacement of the ordinarily used magnetic circuit apparatus by two plates of ordinary soft steel of 1.5 mm thickness. This permitted a sufficiently broad pass band and a small damping time to be obtained.

*This circuit was investigated by Yu. Yu. Sikachevskii.

2. Electromagnetic Contact Frequency Relay

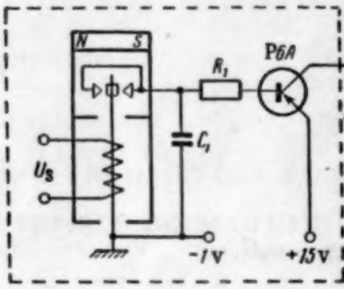


Fig. 4.

The functional schematic of an electro-mechanical contact frequency relay is shown in Fig. 4. The contacts of the electromechanical filter relay are connected to a single-stage semiconductor amplifier, whose load is the winding of a small-size telephone relay (RKM). Capacitor C_1 is used for smoothing the pulsating current in the base circuit, and capacitor C_2 is used for smoothing in the collector circuit and for increasing the relay's operate time (to increase the relay's stability to noise pulses).

In normal operation, the voltage on the frequency relay's contacts is about $U_k \leq 0.8$ volt and the current flowing through the relay contacts is $I_k \leq 0.5$ milliampere. With such voltages and currents, contact erosion due to electrical phenomena does not occur [4, 5]. The working frequency bands cited earlier for the contactless relay are obtained with input signal levels of $U_s = (1 \pm 0.5)U_{N.s}$; the nominal value of signal voltage can be taken as equal to $U_N \approx 5$ volts.

Silver was used as the relay contact materials. The relay's operate band varied relatively little (< 0.5%) in the temperature range from 0 to 40°C.

With the nominal signal voltage $U_{N.s}$, the operate time of the frequency filter together with the RKM relay connected at the output lay within the limits of 0.1 to 0.3 second. Thanks to the retardation of operation of the executive relay, the frequency relay had low sensitivity to noise pulses and shocks. When dc voltage pulses, $U_s \leq 10U_{N.s}$ were applied to the frequency filter's winding, the frequency filter did not operate. The power for reliable relay operation was approximately 5 milliwatts, the input impedance was $R_{in} \geq 3$ kilohms, the dimensions were 30 × 30 × × 60 mm. The bracing of the relay was standard, analogous to the bracing of an RKM telephone relay.

A description of an analogous relay with somewhat different construction is provided in [6].

3. Crystal Triode Frequency Relays

The circuit for a frequency relay realized with plane crystal triodes (Fig. 5) consists of a resonant RC-amplifier with a twin-T bridge in the negative feedback path. As is well known, the operation of a twin-T bridge in amplifier circuits tends to provide a no-load mode and a given voltage since, in this case, the maximum Q factor of the circuit is reached. In order to make the operation of the twin-T bridge approach the no-load mode, the bridge output in the circuit shown was loaded by a triode connected in the circuit with a grounded collector. The output triode, KT_4, to whose collector is connected the winding of an RKM-type relay, is a power amplifier. Its input impedance is low and therefore it is connected to the output of amplifying triode KT_2 via cathode follower KT_3 whose input impedance is slightly shunted by the output of triode KT_2. If a sensitive output relay is employed (polarized, magnetic, etc.), the necessity of triodes KT_3 and KT_4 diminishes. The winding of such an output relay can be connected in the collector circuit of KT_2 via a dividing capacitor and a succeeding rectifying device.

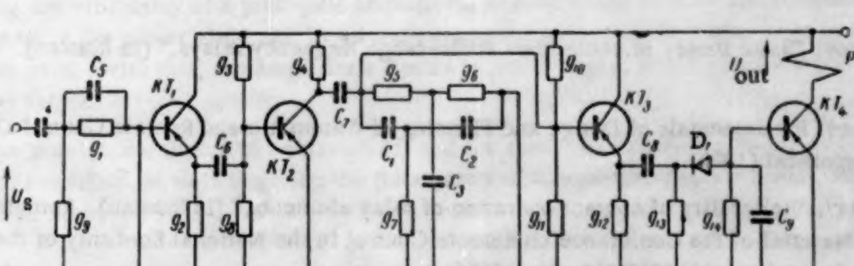


Fig. 5.

Analytical investigations of the circuit showed that its gain K and its Q-factor Q at the quasi-resonant frequency depend in a complicated way on the circuit parameters and the twin-T bridge elements:

$$|\dot{K}_0| = \frac{a r_m g_1 (1+\gamma^2)}{\{[a^2 l r_e H + N(\lambda' H + \gamma\xi) - \xi^2 m]^2 + [(\lambda' H + \gamma\xi) m\xi + N]^2\}^{1/2}};\qquad (3)$$

$$Q = \frac{|\dot{K}_0|\, g_5 \frac{m}{m+1}}{g_1 \sqrt{1+\gamma^2}},\qquad (4)$$

where

$$H = 1+\gamma^2,\qquad N = g_1 H + \gamma m \xi$$

$$a = \frac{1}{r_e(r_k + r_b) + r_b(r_k - r_m)},\ l = r_k - r_m + r_e,\ \xi = g_5\gamma + \omega_0 C,$$

$$\lambda' = g_4 + g_0 + a(r_b + r_e),\ m = \frac{g_5}{g_6},\ n = \frac{C_2}{C_1},\ \gamma = \frac{1/\omega_0 C_1 + 1/\omega_0 C_2}{1/g_5 + 1/g_6},$$

r_e, r_b, r_k and r_m are parameters of the crystal triodes, ω_0 is the quasi-resonant frequency and g_0 is the input conductance of triode KT_3.

For the chosen frequency range (60 to 300 cycles) and m = n, the optimal value of the Q factor is obtained from m = 0.5 to 0.7, γ = 0.7 to 1, g_5 = (1 to 2) × 10^{-5} mho. The remaining parameters of the twin-T bridge are chosen in accordance with the condition that it be tuned to the given quasi-resonant frequency [7]. Experimental data on the relay prepared at the IMA showed the practical independence of the width of the operate band from the temperature in the range 0 to +35°C, adequate stability of the resonant frequency (drift of less than 0.5%) with variations of ±50% of the supply voltage and variations of the magnitude of the input signal, $U_s \leq (1 \pm 0.5) U_{N.s}$. The width of the operate band increased with increased supply voltage and magnitude of input signal.

The relay's input impedance was R_{in} = 40 to 50 kilohms, the operate power was $P_{op} \approx 2$ microwatts. For the nominal value of the input signal voltage, the width of the operate band was from 3 to 5% of the resonant frequency. The relay dimensions were 25 × 35 × 70 mm.

Received April 3, 1959

LITERATURE CITED

[1] N. T. Tynnaya and V. N. Mikhailovskii, "A system of dispatcher control of dispersed objects with frequency selection," [in Russian]. Scientific-technical Conference on Questions of the Development and Production of Apparatus for Remote Control and Communication Channels for Remote Control, convened by the GNTK SM SSSR and NTO Énerg. Prom. Leningradskoo oblasti. (Leningrad, 1958).

[2] N. A. Lifshits, Frequency Selectors [in Russian]. Izd. Voennoi Akademii RKKA im. Budennovo (Leningrad, 1938).

[3] F. O. Katkov, "Basic theory of contactless audio-range frequency relays," [in Russian]. Avtomatika 2 (1956).

[4] B. S. Sotskov, Fundamentals of Design and Planning of Automatic and Remote Control Device Elements, [in Russian]. Gosénergoizdat (1953).

[5] B. S. Sotskov, "Reliability of contact operation of relay elements," [in Russian]. Remote Control in the National Economy. Material of the Conference on Remote Control in the National Economy of the SSSR (Nov. 29 to Dec. 4, 1954). Izdatelstvo AN SSSR (Moscow, 1956).

[6] H. Bopp, "Frequency relays and several possibilities of applying them," [in German]. ETZ 21 (1957).

[7] V. P. Sigorskii and Yu. I. Spektor, "On the analysis of twin-T circuits," [in Russian]. Trudy In-ta mashinovedeniya i avtomatika AN USSR. Questions of Automation and Measurement Technology, 4 (1955).

USING TRANSISTORS TO INCREASE THE EFFICIENCY OF REVERSIBLE DC MAGNETIC AMPLIFIERS

M. A. Rozenblat and G. V. Subbotina

(Moscow)

A new circuit for a push-pull magnetic amplifier with dc load is described, the high efficiency of this circuit being obtained by the use of transistors as switches which permits the harmful effects of the halves of the circuit on each other to be practically eliminated.

The paper provides results of the investigation of a magneto-semiconductor amplifier built in accordance with the high-efficiency circuit presented here, and the data are also given for a high-stability measuring amplifier constructed on the same basis.

The well-known push-pull magnetic amplifiers with dc loads are distinguished by their common disadvantage — low efficiencies, not exceeding 50% of that in the most effective circuits, even when there is no loss in the windings, rectifiers or cores. The reason for this phenomenon comes down to the shunting interaction of the two halves of the push-pull circuits which leads to an undesirable cross distribution of currents. To overcome this limitation one connects a ballast resistor at the amplifier output, and 50 to 83% of the output power is lost in this resistor [1].

New possibilities for improving the characteristics of magnetic amplifiers, push-pull amplifiers in particular, are revealed by the joint usage of them and transistors [2-5].

In the present paper we give the results of the development and investigation of a push-pull magnetic amplifier based on a new circuit.* In this circuit, the transistors act as switches which, in dependence on the sign of the signal, permit the load to be switched to the output of that one of the two single-ended amplifiers whose output current is increasing. With this we succeeded in almost completely using the power of the single-ended amplifiers when they operate in the push-pull circuit with a single dc load.

Increasing the efficiency of a push-pull amplifier is accompanied both by an increase in the power yield and by an increase in the power yield and by an increase in the extent of the linear portion of its characteristic and in its power gain. With this, the amplifier's inertia is not changed, which gives a significant increase in its dynamic energy factor.

Below, we provide the theory of operation of, and the results of experimental investigation of, a magneto-semiconductor amplifier, as well as giving the parameters of a high-stability measuring amplifier based on the circuit considered here.

Circuit and Theory of Operation of the Amplifier

The circuit of the magneto-semiconductor amplifier (Fig. 1) is an ordinary differential push-pull magnetic amplifier with internal feedback in which the rectified current from the single-ended arms is applied to the load via emitter-collector junction transistors T1 and T2.

*M. A. Rozenblat. Magneto-semiconductor amplifier [in Russian]. Patent application (Avtorskaya zayavka) 623120/26.

The special feature of the circuit is that there is applied to the triodes, which are connected by a common emitter, a controlling voltage which is proportional to the output voltage of the corresponding single-ended magnetic amplifier. The magnitude of the controlling current is established by a divider made up of resistors R_1 and R_b.

With the application of a control current of a definite sign to the magnetic amplifier's input, a large current flows from one arm to the load, generating a significant voltage between the emitter and base of the corresponding triode and opening it, while the base current of the second triode is sharply dropped and the degree of opening of this triode is decreased.

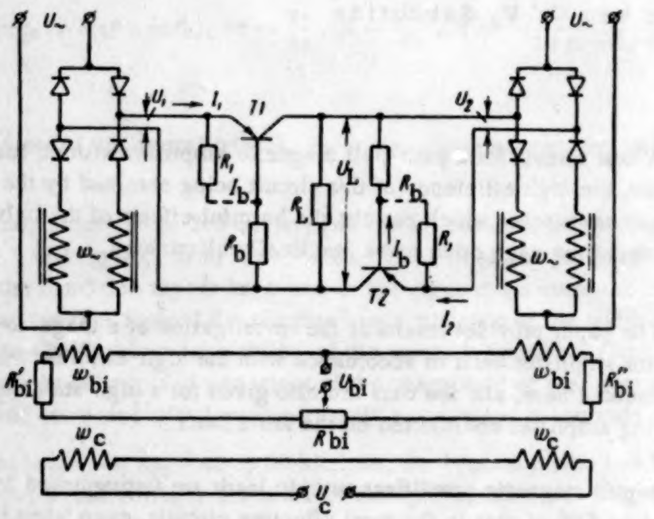


Fig. 1.

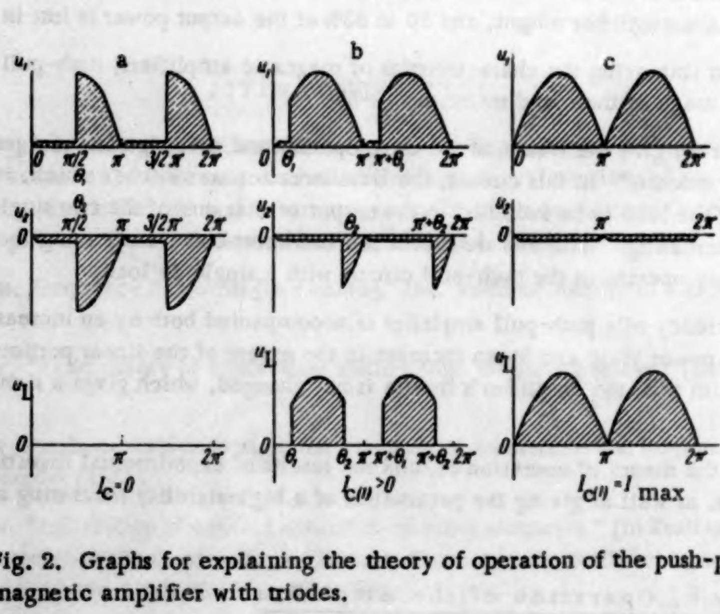


Fig. 2. Graphs for explaining the theory of operation of the push-pull magnetic amplifier with triodes.

Thus, the second triode will present a large impedance to the first arm's current which both limits the harmful current shunting and increases the effectiveness of the entire circuit's operation. Then the signal current changes, the roles of the triodes are changed and a current flows in the load in the opposite direction with the same result as far as circuit effectiveness is concerned.

If, with minimum current in an arm, the corresponding triode is completely cut off, then the maximum current in the load of the push-pull amplifier will equal the maximum current of its single-ended arm.

A graphic presentation of the operation of the amplifier can be obtained if we consider the graphs of the output voltage of a magnetic amplifier whose cores have rectangular hysteresis loops. We assume, for simplicity, that in the absence of a signal the saturation angles of both single-ended amplifiers is $\theta_1 = \theta_2 = \pi/2$, thereby obtaining the curves of the output voltages U_1 and U_2 which are given in Fig. 2,a. In this case, the magnetic amplifier's output voltage U_L equals zero, independently of the presence of the triodes.

When a signal is applied to the amplifier's input, the saturation angle θ_1 of the first single-ended amplifier decreases while that of the second, θ_2, increases by the same amount (Fig. 2,b). In this case, during the interval $0 \leq \omega t \leq \theta_1$ there are, at the outputs of both amplifiers, small, practically identical, free-running voltages which, with the proper choice of triode mode of operation, are insufficient to open the triodes. Initial triode collector current I_{c0} flows through the emitter-collector junctions, and the resulting load voltage (current) equals zero. In the interval $\theta_1 \leq \omega t \leq \theta_2$ a voltage is applied to the input of the first triode (in the emitter-base circuit) which suffices to open the triode, while the second triode remains closed as before. Therefore, current I_1 at the first amplifier's output is, for all practical purposes, not shunted through the second amplifier's rectifier, but falls completely on the load.

In the third interval, $\theta_2 \leq \omega t \leq \pi$, after the core of the second amplifier has been saturated, both triodes conduct. Since the output voltages (currents) have identical instantaneous values, the load voltage in this interval will be zero. Therefore, the output voltage U_L (or current) of the amplifier has the form shown in the lower graphs of Fig. 2,b. The forms of the curves for voltages U_1, U_2 and U_L with maximum signal at the amplifier's input are shown in Fig. 2,c. In this case, the second triode is cut off during all the periods and the load current is virtually equal to the current at the output of one single-ended amplifier.

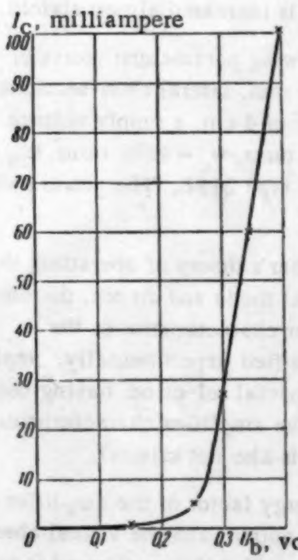


Fig. 3. The function $I_c = f(U_b)$ for transistor 313B.

In an actual amplifier the divider parameters R_1 and R_b must be so chosen than the triodes do not limit the magnitude of the current at the outputs of the single-ended amplifiers in the interval of saturated cores, $\theta \leq \omega t \leq$. This condition will be met if, for $\theta \leq \omega t \leq \pi$, the collector current is

$$i_c = I_{c0} + \frac{\alpha}{1-\alpha} i_c > i,$$

where $\underline{i}$ is the instantaneous value of the current at the output of a single-ended magnetic amplifier, I_{c0} is the initial collector current, i_b is the base current corresponding to amplifier output current $\underline{i}$ and established by means of resistors R_1 and R_b, α is the current gain of the triodes with a common base. It should however, be borne in mind that, in a single-ended circuit, the load current is $i_L = i - i_b$. Therefore the choice of the triodes' mode of operation, as determined by R_1 and R_b, must also guarantee, for maximum efficiency, that the value of base current i_b is much less than the current $\underline{i}$. In practice; it suffices if $i_b < 0.1\underline{i}$. Implementing these two conditions is easily done, since the current gain of triodes with common emitters, $[\alpha/(1-\alpha)]$, is ordinarily not less than 15 to 20 and, in the best types, is considerably greater.

The presence of a magnetizing current (voltage) at the output of the second single-ended amplifier in the interval $\theta_1 \leq \omega t \leq \theta_2$ can somewhat open the corresponding triode, which leads to a certain lowering of the load current of the push-pull amplifier. Ordinarily this effect is negligibly small due to the nonlinear character of the function $I_c = f(U_b)$ in the region of small values of U_b (Fig. 3).

It is clear from the theory of operation we have been considering that the choice of parameters and design of the push-pull magnetic amplifier reduces, essentially, to the choice of parameters and design of the corresponding single-ended magnetic amplifier. If the triode parameters are properly chosen they have practically no effect

on the operation of the single-ended amplifier. It is not difficult to choose the necessary type of triode once the amplifier's output parameters are given. The triode's maximum current equals the maximum current of the single-ended amplifier. The maximum voltage on the triode occurs when the output current of the corresponding single-ended magnetic amplifier is a minimum, and equals its maximum output voltage. Therefore, the triodes are chosen in accordance with the conditions that $I_{c\ adm} \geq I_{1\,max}$ and $U_{c\ adm} \geq U_{1\,max}$.

We note that the suggested method of connecting triodes is completely applicable to other push-pull amplifier circuits with dc outputs, including circuits with external feedback paths, used for increasing zero stability [1], for center point rectification circuits, etc.

EXPERIMENTAL DATA

Experimental investigation of a number of amplifiers based on the circuit of Fig. 1 showed that the push-pull amplifier's maximum load current is less than the current at the output of the corresponding single-ended amplifier by not more than 5-to 10%, which was due to the partial shunting of current i_1 through the triode's emitter-base circuit and to the effect, cited above, of the free-running current of the second single-ended amplifier (Fig. 4, b and c). We note that a further approximation of the push-pull amplifier's load current to the load current of the single-ended amplifier can be attained by using a nonlinear impedance for R_b (for example, a thyrite type). In this case, a decrease of voltage at the single-ended amplifier's output increases the magnitude of impedance R_b which leads to an additional limitation of the current through the given triode. The use of nonlinear impedances and of other well-known means for decreasing magnetic amplifiers' free-running current are advantageous in cases when the single-ended amplifiers have a small range of load current variation.

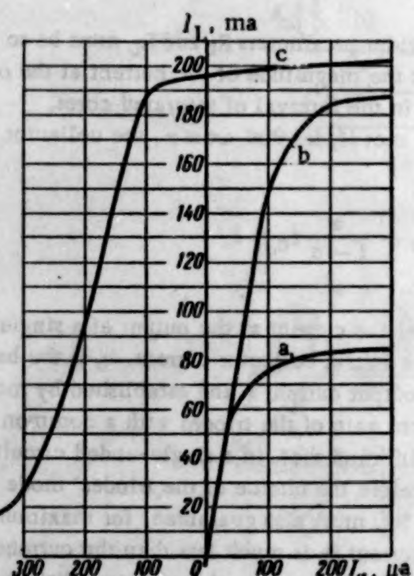


Fig. 4. Magnetic amplifier characteristics: a is for a push-pull circuit with ballast resistors, b is for a push-pull circuit with triodes and c is for a single-ended amplifier.

Figure 4 shows the characteristics of one and the same push-pull amplifier: a is for an ordinary circuit with ballast resistors and b is for the suggested circuit, with the ballast resistors replaced by triodes. In the second case, with identical load impedance, the maximum load current is increased by a factor of approximately 2.4 and the power is increased almost sixfold.

The amplifier has the following parameters: cores of alloy 65NP with thickness of 0.15 mm, lateral cross-sectional area of 1 cm² and mean diameter of 4 cm, a supply voltage of 9 volts at 50 cycles, $w_{\sim}$ = 300 turns, w_c = 3500 turns, R_{in} = = 600 ohms, R_L = 20 ohms, triode type 313B. The power gain of the circuit of Fig. 1 was 100,000.

It is clear, from the amplifier's theory of operation, that the triodes operate in a noncritical mode and do not, therefore, have much effect on the amplifier characteristics or the stability of the null. This was verified experimentally. Replacement of the triodes without any special selection having been made had virtually no effect on the amplifier characteristics. The choice of resistors R_1 and R_b is also not critical.

The increased dynamic energy factor of the amplifier due to the decreased power loss at its output, and the virtual absence of triode effect on the amplifier's null stability allowed it to be used for implementing a dc measuring amplifier. For this purpose, a negative electric feedback path was connected from the amplifier's output to its input. With this, the positive feedback path was increased by several turns for the purpose of increasing the open-loop gain. By doing this we were able, with a single-ended amplifier whose parameters were given above, to obtain, with no special selection of circuit elements and with a gain of 2000, a constancy of gain to within 0.2% when the supply voltage varied by 10 to 15% and the load impedance varied by a factor of three. With this, the amplifier's time constant equalled three periods of supply voltage. The null drift of the amplifier over an 8-hour period of continuous operation corresponded to a signal power of 1.5×10^{-10} watts. With a stabilized voltage supply, the accuracy of the amplifier's operation increases.

SUMMARY

The method suggested for increasing the efficiency and the energy factor of push-pull magnetic amplifiers with dc outputs is based on the use of transistors operating as controlling switches. With this, the triode (transistors) parameters are not critical and, if properly chosen, have little effect on the single-ended amplifier characteristics.

The suggested method of constructing push-pull dc magnetic amplifiers can also be used both for low-power high-sensitivity amplifiers and for power output stages. The maximum power of such amplifiers is determined by the parameters of the triodes.

Received January 8, 1959

LITERATURE CITED

[1] M. A. Rozenblat., Magnetic Amplifiers [in Russian]. Sovetskoe Radio (1956).

[2] V. S. Volodin, E. D. Larin, M. A. Rozenblat and G. V. Subbotina, "High-speed push-pull magneto-semiconductor amplifiers for power servo drives," [in Russian]. Automation and Remote Control (USSR) 20, 3 (1959).*

[3] R. L. Bright, Junction transistors used as switches, Trans. AIEE, vol. 74, pt. I, (1955).

[4] A. G. Milnes, High-efficiency push-pull magnetic amplifiers with transistors as switched rectifiers, Commun. and Electronics, No. 37, 1958.

[5] O. A. Kossov, "Magneto-crystal amplifiers," [in Russian] Automation and Remote Control (USSR) 20, 7 (1959). *

* See English translation.

INVESTIGATION OF THE OPERATION OF A MAGNETIC AMPLIFIER WITH SELF-SATURATION ON A THREE-PHASE LOAD

A. L. Pisarev

(Moscow)

The steady-state operation is considered for a magnetic amplifier with self-saturation operating on an active three-phase load for amplifiers with free and with suppressed ac components in the control windings.

Methods are given for constructing the control characteristics, and the question of the voltage on the rectifiers is considered for both cases.

The use of magnetic amplifiers with self-saturation for controlling three-phase loads has a number of advantages as compared with the use of ordinary saturated coils.

The basic advantage is the significant increase in gain which, in a number of cases, allows one to do without intermediate amplifiers, and to reduce significantly the size and weight of the entire device. The presence of rectifiers in the load circuit complicates the device but, as will be shown below, the quantity of these can be small.

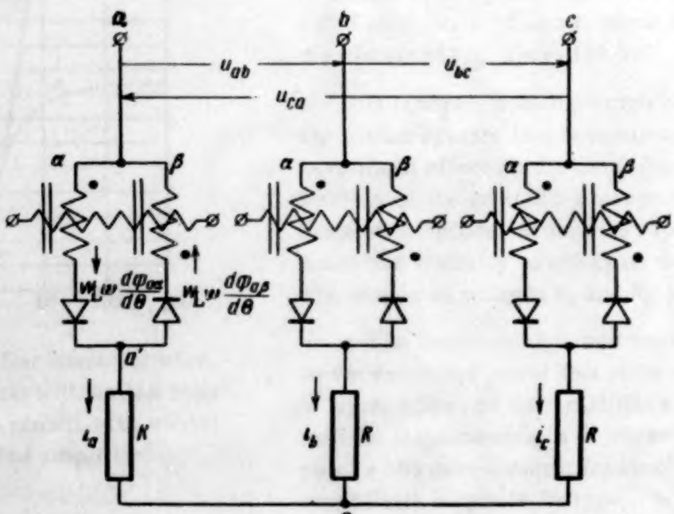


Fig. 1.

The circuit of a magnetic amplifier with self-saturation and a three-phase load is shown in Fig. 1.

For the analysis of the circuit we make the following assumptions.

1. The circuit is acted upon by a symmetric three-phase system of sinusoidal voltages.

2. The amplifier has an active load.

3. The rectifiers have a negligibly small forward impedance and an infinite back impedance.

4. The dynamic hysteresis loop of the core material is approximated by the parallelogram with the parameters shown in Fig. 2.

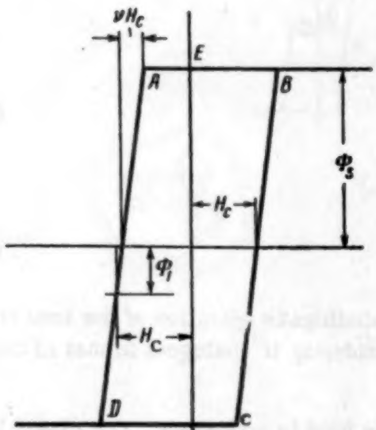


Fig. 2.

5. The active impedances of the amplifier's load windings are assumed to be negligibly small.

6. The supply voltage (phase) is chosen in accordance with the relationship

$$U_m = \frac{2}{\sqrt{3}} \Phi_s W_L \, \omega \cdot 10^{-8}. \tag{1}$$

The basis of expression (1) are given in [1].

7. The magnetizing current is negligibly small in comparison with the current in the saturated portion of the period.

It is known, from amplistat theory [2], that their properties depend essentially on whether or not ac components can be found in the control winding circuit. In three-phase circuits, just as in single-phase ones, two extreme modes of amplifier operation can be distinguished: operation with free ac components in the control winding and operation with suppressed ac components. A magnetic amplifier (Fig. 1) contains six reactors, two per phase. Each reactor is denoted by a double subscript. For example, the subscript $a\alpha$ denotes reactor α phase $\underline{a}$, etc. The reactor control windings may be either individual for each reactor or common to each pair of reactors (as shown on Fig. 1).

If the control windings of all phases are connected in parallel to a source of control voltage with low internal impedance, then the mode of amplifier operation with free ac components in the control windings will occur.

If the control windings are connected in series to a control voltage source with high internal impedance then the mode of amplifier operation with suppressed ac components in the control windings will occur. It should be mentioned that, when the phase control windings are connected in series, even with a very low impedance in the control circuit, the magnetic amplifier's mode of operation will approximate to the mode with suppressed ac components in the control windings, since the even current harmonics, not multiples of three (second, fourth, etc.), cannot flow in the control circuit (their sum is zero).

We shall consider further these two modes of magnetic amplifier operation.

1. Amplifier with Free ac Components in the Control Windings

Similarly to single-phase systems, the following conditions will be valid for each pair of reactors.

1. Due to the presence of a very low-impedance control winding-loop shunting the pair of reactors, the variations in the reactors' magnetic flux will occur simultaneously and with identical velocities. For the reactors of phase $\underline{a}$, for example, the following relationship is valid:

$$\frac{d\Phi_{a\alpha}}{d\theta} = -\frac{d\Phi_{a\beta}}{d\theta}. \tag{2}$$

2. When one of the phase reactors is saturated, the magnetic flux of the other reactor maintains an invariable value

$$\Phi_{a\beta} = \text{const for } \Phi_{a\alpha} = \Phi_s. \tag{3}$$

In [1] there was considered the operation on a three-phase load of a magnetic amplifier with series-connected load windings. It was shown that, for an amplifier with free ac components in the control winding, three modes of operation of the load circuit could occur, depending on the size of the saturation angle θ_s, these modes being differentiated by the form of the load current curve and by the characteristic equation of the circuit's electrical equilibrium. The boundaries of these modes are defined by the following values of the saturation angle:

for the first mode

$$\frac{\pi}{2} < \theta_s < \frac{5\pi}{6}; \tag{4}$$

for the second mode

$$\frac{\pi}{3} < \theta_s < \frac{\pi}{2}; \tag{5}$$

for the third mode

$$0 < \theta_s < \frac{\pi}{3}. \tag{6}$$

We now show that when there are free ac components in the control windings, the operation of the load circuit of the magnetic amplifier with self-saturation which we have been considering is analogous to that of the amplifier without self-saturation which was considered in [1].

In fact, the two conditions cited above on the magnetic flux variations hold in both cases. The closed loop of the control windings, in both cases, guarantees that equality (2) holds. The difference amounts to this, that in the interval of magnetic flux variation the network voltage is divided equally between the two reactors of a given phase in the magnetic amplifier with series-connected load windings while, in the analogous interval for the given case, all the net voltage is applied to each of the reactors.

Moreover, when one of the phase reactors is saturated in the magnetic amplifier with series-connected load windings the load current flows in the load windings of both reactors, while in the analogous interval of the given case, the load current flows only in the winding of the saturated reactor. With this, the load circuit of the second reactor is cut off by a rectifier. In the given case, the reactors operate alternately.

These special features however, do not disturb the analogy of the reactors' operation as electrical circuit elements, since the same three modes of load circuit operation occur in the case under consideration as in the case of a circuit with series-connected load windings. The form of the curve for current in the load is identical for both circuits. Figure 3,a shows the curve of the load current for phase a for the case corresponding to the mode described by (5). This same curve can be seen on the oscillogram of Fig. 4.

Based on the analogies of the physical processes in the reactors, the equations for the electrical equilibrium of the given circuit, set up in segments defined by the magnetic states of the reactors, are identical with the equations which define the operation of a magnetic amplifier with series-connected load windings in the analogous modes [1]. The only difference is that, in the equations defining the increments of the reactors' magnetic fluxes, the denominators will contain W_L instead of $2W_L$. The validity of this substitution is obvious and requires no explanation.

Figure 3,b shows the curves of the magnetic fluxes of reactors $a\alpha$ and $a\beta$. The expressions for these magnetic fluxes are defined as follows. In the interval $\theta_{sa\alpha} - \frac{\pi}{3} < \theta < \theta_{sa\alpha}$ (Fig. 3) reactors $b\beta$ and $c\alpha$ are saturated and the magnetic flux of reactors $b\alpha$ and $c\beta$ are invariant. The magnetic flux of reactor $a\alpha$ increases and that of reactor $a\beta$ decreases.

The following equations are valid:

$$u_{bc} - (i_c - i_b)R = 0, \tag{7}$$

$$u_{ca} - e_{a\alpha} + i_c R = 0, \tag{8}$$

$$i_c = -i_b. \tag{9}$$

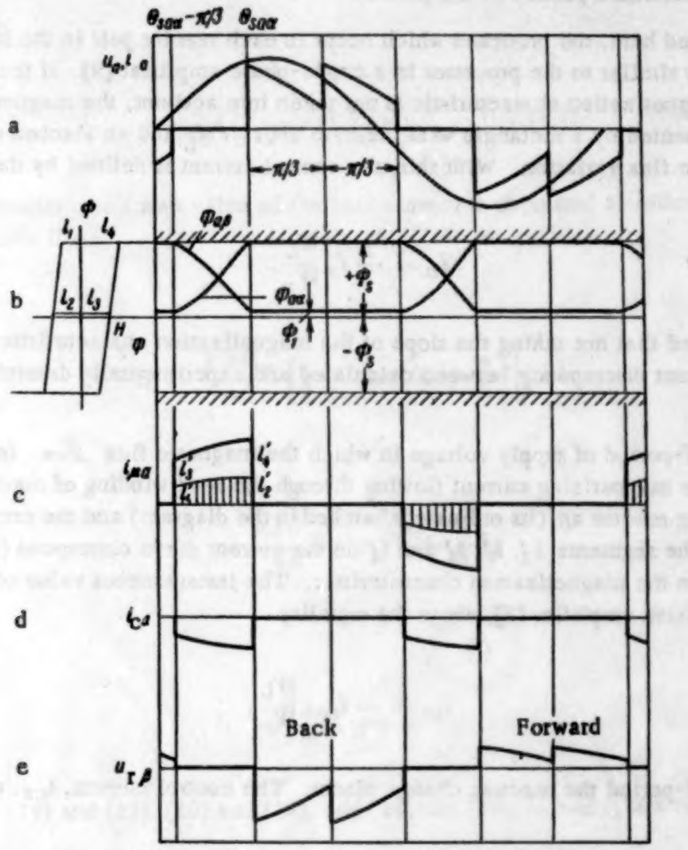


Fig. 3. Relationships for the magnetic amplifier with free ac components in the control windings.

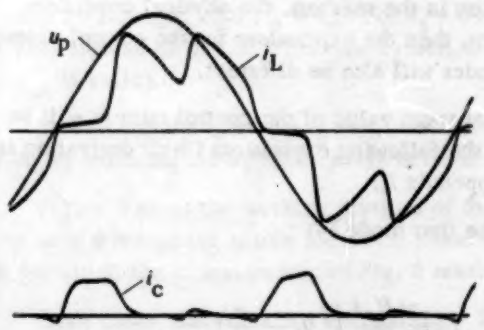


Fig. 4. Load and control current oscillograms for a magnetic amplifier with free ac components in the control windings.

The last two equations are correct if assumption 7 is taken into account.

We find from equations (7), (8) and (9) that

$$e_{a\alpha} = u_{c\alpha} + \frac{u_{bc}}{2}. \tag{10}$$

For $\theta = \theta_{sa\alpha} - \frac{\pi}{3}$ $\varphi_{a\alpha} = \Phi_1$, $\varphi_{a\beta} = \Phi_s$, and therefore the expressions for the instantaneous value of magnetic flux will be

$$\varphi_{a\alpha} = \Phi_1 + \frac{10^8}{W_L\,\omega} \int\limits_{\theta_{sa\alpha} - \frac{\pi}{3}}^{\theta} \left(u_{c\alpha} + \frac{u_{bc}}{2} \right) d\theta = \Phi_1 + \frac{10^8}{W_L\,\omega} \int\limits_{\theta_{sa\alpha} - \frac{\pi}{3}}^{\theta} 1.5\, u_a\, d\theta, \tag{11}$$

$$\varphi_{a\beta} = \Phi_s - \frac{10^8}{W_L\,\omega} \int\limits_{\theta_{sa\alpha} - \frac{\pi}{3}}^{\theta} 1.5\, u_a\, d\theta. \tag{12}$$

Figure 3,c represents the magnetizing component of the load current of phase a (in a scale different from that of the current in the saturated portion of the period).

In the case considered here, the processes which occur in each reactor pair in the interval of magnetic flux variation are qualitatively similar to the processes in a single-phase amplistat [2]. If the slope of the lateral sides of the material's magnetization characteristic is not taken into account, the magnetizing component of the load current will be represented by a rectangle with ordinate $2H_C l_C/W_L$ and an abscissa determined by the duration of the interval of magnetic flux variation. With this, the control current is defined by the expression

$$I_C = \frac{1}{2} I_\mu \frac{W_L}{W_y}.$$

However, tests showed that not taking the slope of the magnetization characteristic's lateral sides into account leads to a significant discrepancy between calculated and experimentally determined amplifier control characteristics.

We consider the half-period of supply voltage in which the magnetic flux $\Phi_{a\alpha}$ increases while magnetic flux $\Phi_{a\beta}$ decreases. The magnetizing current flowing through the load winding of reactor $a\alpha$ (Fig. 3,c) includes the current for magnetizing reactor $a\beta$ (its ordinate is hatched in the diagram) and the current for magnetizing reactor $a\alpha$ itself. With this, the segments l_1', l_2', l_3' and l_4' on the current curve correspond (in a changed scale) to the segments l_1, l_2, l_3 and l_4 on the magnetization characteristic. The instantaneous value of the control current in this case, as for a single-phase amplifier [3], obeys the equality

$$i_{Ca} = -i_{\mu a\beta} \frac{W_L}{W_C}. \tag{13}$$

In the following half-period the reactors change places. The control current, i_{Ca}, of phase a is shown in Fig. 3,d.

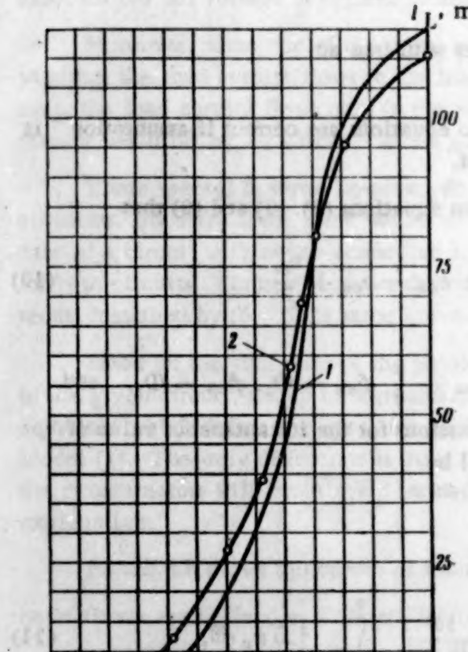


Fig. 5. Computed and experimentally determined control characteristics of a magnetic amplifier with free ac components in the control windings.

For illustration, Fig. 4 gives the oscillograms of the phase voltage u_p, the load phase current i_L and the phase control current i_C. Since, in the three modes defined by (4), (5) and (6), in the interval of magnetic flux variation in the reactors, the physical conditions are different, then the expressions for the control current in these modes will also be different.

The average value of the control current will be defined by the following expressions (their derivation is given in Appendix I):

for the first mode (4)

$$I_C = \frac{-H_c l_c}{\pi W_C}\Big[\Big(\theta_s - \frac{\pi}{6}\Big) - \tag{14}$$
$$- \nu(-1.05 + 1.38 \sin\theta_s - 0.4\cos\theta_s)\Big];$$

for the second mode (5)

$$I_C = \frac{-H_c l_c}{\pi W_y}\Big[\frac{\pi}{3} - \nu\Big(\frac{\pi}{3} + 0.6\cos\theta_s - 0.7\sin\theta_s\Big)\Big]; \tag{15}$$

for the third mode (6)

$$I_C = \frac{-H_c l_c}{\pi W_c} [\theta_s - \nu(1.73 \sin \theta_s - 0.73\,\theta_s)]. \qquad (16)$$

For the enumerated modes, the mean value of the load current is expressed as follows (the derivation of the formulae is given in Appendix II):

for the mode of (4)

$$I_L = \frac{\sqrt{3}\,U_m}{\pi R}\left[1 + \cos\left(\theta_s + \frac{\pi}{6}\right)\right]; \qquad (17)$$

for the mode of (5)

$$I_L = \frac{\sqrt{3}\,U_m}{\pi R} \sin\left(\theta_s + \frac{\pi}{3}\right); \qquad (18)$$

for the mode of (6)

$$I_L = \frac{U_m}{\pi R}(1 + \cos\theta_s). \qquad (19)$$

Thus, in expressions (14) and (17), (15) and (18), and (16) and (19), I_C and I_L are related to each other via the parameter θ_s.

Computation of the magnetic amplifier's control characteristics is carried out by determining the values of I_C and I_L for a number of given magnitudes of θ_s.

When the magnetizing component of the load current must be taken into account, this can be done approximately by assuming that $I_\mu \equiv I_C$. For $\theta_s = 5\pi/6$, it is easily seen than $I_\mu = 2I_C W_C / W_L$ (since with this $\Delta\Phi = 2\Phi_s$ and, instead of partial hysteresis cycles, there occurs a symmetric cycle and, consequently, the equation $i_{\mu a\alpha} = 2i_{\mu a\beta}$ is valid).

By drawing the quantity I_μ (segment AB on Fig. 50) and then joining point B with point O, we obtain the line BO whose ordinates are approximately defined by the magnetizing component of the load current I_μ.

Figure 5 gives the working portions of the computed control characteristic, 1, for a magnetic amplifier operating on a three-phase active load with phase control windings in parallel. Below is given the data for the amplifier for which the characteristic of Fig. 5 resulted.

Core data: the cores are toroidal; $l_C = 11.55$ cm; $q_C = 0.685$ cm^2; core material was 65NP; dynamic coercive force was $H_C = 0.26$ ampere-turns per centimeter; $\nu = 0.27$. Circuit data: $R = 285$ ohms; $U_p = 29$ volts; $W_L = 1500$; $W_C = 200$. The rectifiers were type DG-Ts24.

Figure 5 also gives the experimentally determined control characteristics (curve 2).

2. Amplifier with Suppressed ac Components in the Control Windings

As was already stated, this mode of magnetic amplifier operation can occur when the control windings of the individual phases are connected in parallel and supplied from a source with a significant internal impedance.

We consider the processes occurring in the reactors of one of the phases (for example, phase a). With the assumptions that were made regarding the smallness of the load windings' active impedance and the rectifiers' back impedance, we can easily show that, in this case, conditions (2) and (3) also remain valid.

A pair of reactors of one of the phases can be likened to a single-phase magnetic amplifier with self-saturation and suppressed ac components in the control winding, the properties of which are very well known [2, 4, 5]. If, at some moment of time, reactor aβ ceases to be saturated then, starting with this moment, the reactor begins to act like an inductance. The current in its winding retains its previous direction. For this interval, the following equation is valid:

$$u_{a0} - W_L\omega \frac{d\phi_{a\beta}}{d\theta} - i_a R = 0. \tag{20}$$

Here, u_{a0} is the voltage between points a and 0 (Fig. 1).

Part of voltage u_{a0}, equal to $u_{aa'} = W_L\omega \frac{d\phi_{a\beta}}{d\theta}$, is also applied to reactor aα, forcing magnetic flux $\phi_{a\alpha}$ to increase and rectifier aα to begin switching. For the load windings loop, if the active impedance of this loop is neglected, the following expression will be valid

$$W_L\omega \frac{d\phi_{a\beta}}{d\theta} + W_L\omega \frac{d\phi_{a\alpha}}{d\theta} = 0 \tag{21}$$

where

$$\frac{d\phi_{a\alpha}}{d\theta} = -\frac{d\phi_{a\beta}}{d\theta}.$$

Equality (2) holds during the entire interval of reactor flux variation and, therefore, the intervals of magnetic flux variation of both reactors coincide in time.

When reactor aα is saturated, the load winding of reactor aβ is short-circuited and, consequently, $\phi_{a\beta}$ = = const. It thus turns out that, with the assumptions made, conditions (2) and (3) are met in magnetic amplifiers with either free or suppressed ac components in the control windings. The difference amounts to this, that in the first case the meeting of the condition on the smallness of the load windings' active impedance and the rectifiers' back impedance is optional, since the closed loop of control windings of each pair of reactors necessarily guarantees that conditions (2) and (3) will hold. In the second case, it is mandatory that this condition be met.

In practice, this condition is met in relatively powerful magnetic amplifiers, where R >> r and r is the active impedance of the load winding. The case of finite magnitudes of r and r_r, where r_r is the rectifier's forward impedance, will be considered qualitatively below. It is obvious that when conditions (2) and (3) are met, the processes occurring in the magnetic amplifier's load circuit in the given case will be completely identical with those in the amplifier with free ac components in the control windings.

The relationships between the average value of the load current and the saturation angle, (17), (18) and (19), remain valid.

We turn now to the determination of the relationship between the control field strength H_{c0} and the magnitude of the reactor's magnetic flux. It is known from the theory of single-core magnetic amplifiers [5] that if a fall in the reactor's magnetic flux occurs under the action of an external voltage (with switching rectifiers) then the relationship between the instantaneous values of the resulting magnetizing force and the reactor's magnetic flux is given by the descending arm of the dynamic hysteresis loop of the reactor (EAD on Fig. 2). In this case, the difference between the three-phase and the single-phase amplifiers reduces to the difference in the laws of variation of the external voltage applied to the reactors. This circumstance does not affect the character of the relationship between the resulting field strength and the magnetic flux.

It was shown above that, with the assumptions made, during the entire interval of change of the reactors' magnetic flux, these changes occur under the action of the external voltage u_{a0} [Cf., equation (20)] and therefore there can be no additional fall of the magnetic flux after the cessation of current flow in the load circuit. At the moment when the current flow in the load circuit ceases, the resulting reactor field strength becomes equal to H_{c0} and the reactor's magnetic flux becomes equal to Φ_1. In the case considered, therefore, the relationship

between H_{C0} and Φ_1 is defined by the descending arm of the reactor's dynamic hysteresis loop.

The relationship between the magnetic flux increment, $\Delta\Phi = \Phi_s - \Phi_1$, and the saturation angle can be found for the three modes of amplifier operation defined by (4), (5) and (6) by integrating the expressions for the voltages applied to the reactor (Cf., Appendix III). These relationships have the following forms:

for the mode defined by (4)

$$\Delta\Phi = \frac{\sqrt{3}U_m \cdot 10^8}{2W_L\omega}\left[1 - \cos\left(\theta_s + \frac{\pi}{6}\right)\right]; \tag{22}$$

for the mode defined by (5)

$$\Delta\Phi = -\frac{1.5U_m \cdot 10^8}{W_L\omega}\cos\left(\theta_s + \frac{\pi}{3}\right); \tag{23}$$

for the mode defined by (6)

$$\Delta\Phi = \frac{1.5U_m \cdot 10^8}{W_L\omega}(1 - \cos\theta_s). \tag{24}$$

To construct the amplifier's control characteristics, it is necessary to give a number of values of θ_s, to determine for them the corresponding values of load current by use of formulae (17), (18) and (19), and then to determine the magnitude of $\Delta\Phi$ by formulae (22), (23) and (24). Then, knowing the configuration of the limiting dynamic hysteresis loop for the core material (or an approximation to it), one determines the corresponding values of H_{C0} and I_C from the computed values of $\Delta\Phi$.

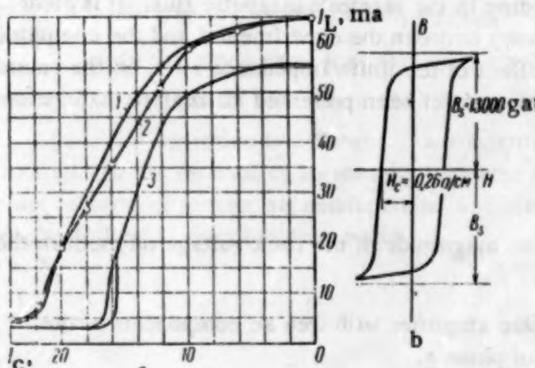


Fig. 6. Characteristics of a magnetic amplifier with suppressed ac components in the control windings.

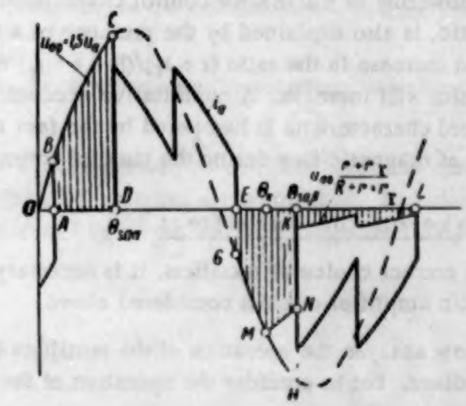


Fig. 7. Relationships for a magnetic amplifier with suppressed ac components in the control windings.

As an example of this construction, Fig. 6 gives the control characteristics, curve 1, of the magnetic amplifier considered in the previous example. In this example, the load impedance is 400 ohms. The experimentally determined characteristic, given by curve 2, is shown for purposes of comparison. Figure 6,b gives the limiting dynamic hysteresis loop for material 65NP which was used in the construction of the control characteristics. With lower values of source voltage the magnetic amplifier's control characteristic has an instability in the region of small load currents, the nature of this instability being the same as in the case of a single-phase amplifier [4]. To illustrate this phenomenon, curve 3 of Fig. 6,a gives the control characteristic of the same amplifier with U_p = 25 volts.

It is clear from Fig. 6,a that the experimentally determined control characteristics are somewhat lower than the computed ones, particularly in the upper portion of the operating segment. This is explained by the fact that, in the construction of the calculated characteristic, an idealization was assumed, amounting to the assumption that $r + r_r = 0$. As was shown above, this assumption excludes the possibility of a fall in reactor magnetic flux after cessation of rectifier switching. In actual amplifiers such a fall always occurs. This may be shown by the same methods as are used for single-phase amplifiers [4]. As an example, we consider the second mode of operation, defined by (5). For any magnetic amplifier in steady-state operation, the area bounded by the curve $W_L\omega\, d\phi/d\theta$ in the interval of increasing magnetic flux equals the area bounded by this curve in the interval of falling magnetic flux. It was shown above [Formula (11)] that in the interval of reactor magnetic flux variation in the mode defined by (5), the voltage

$$u_{a0} = 1.5\, u_a \approx W_L\, \omega \frac{d\phi_{a\alpha}}{d\theta}.$$

On Fig. 7, hatched area ABCD corresponds to the growth of magnetic flux $\phi_{a\alpha}$. If we set $r + r_r = 0$, the fall of magnetic flux $\phi_{a\alpha}$ will correspond to area EGHK. Starting at the moment $\theta_{sa\beta}$ and until the end of the interval of reactor $a\beta$ conductance, $u_{aa'} = 0$. Magnetic flux $\phi_{a\alpha}$ cannot vary. If, however, we assume that $r + r_r \neq 0$, then magnetic flux $\phi_{a\alpha}$ can fall in this interval, since $u_{aa'} = u_{a0}(r + r_r)/(R + r + r_r)$. This voltage must balance the emf of reactor $a\alpha$, $e_{a\alpha} = -W_L\omega d\phi_{a\alpha}/d\theta$. The additional fall of magnetic flux $\phi_{a\beta}$ is denoted by the hatched area in the interval KL.

Thus, if it is assumed that in the entire interval from point E to $\theta_{sa\beta}$ the fall of magnetic flux $\phi_{a\alpha}$ occurs under the action of an external voltage (as was done for the idealized case) then, since $\theta_{sa\beta} = \theta_{sa\alpha} + \pi$, it is found that magnetic flux $\phi_{a\alpha}$ falls by a larger amount than it increases (due to the hatched area in the interval KL), which is impossible. Consequently, one can only assume that magnetic flux $\phi_{a\alpha}$ falls in the interval $E\theta_{sa\beta}$ in accordance with a different law than it increases in the interval $A\theta_{sa\alpha}$. Indeed, at the point θ_k, where the relationship $W_L\omega d\phi_{a\alpha}/d\theta = u_{a0}$ becomes valid (current i_a passes through zero), rectifier $a\alpha$ stops switching and magnetic flux $\phi_{a\alpha}$ begins to fall under the action of the magnetizing force (mf) of the control winding along line MN. With this, obviously, area MHN equals the hashed area in the interval KL.

The lowering of the reactor control characteristic obtained experimentally, as compared with the computed characteristic, is also explained by the presence of a relaxation drop in the reactor's magnetic flux. It is clear that with an increase in the ratio $(r + r_r)/(R + r + r_r)$ the discrepancy between the experimental and the computed characteristics will increase. A quantitative accounting of the effect of the finite impedance $r + r_r$ on the course of the control characteristic is hampered by the fact that there has not yet been presented an analytic expression for the fall of magnetic flux during the time of action of an mf.

3. Voltage on the Rectifiers

For a correct choice of rectifiers, it is necessary to know the magnitude of the back voltage on them in the two magnetic amplifier circuits considered above.

We now analyze the operation of the rectifiers in a magnetic amplifier with free ac components in the control windings. Let us consider the operation of the rectifiers of phase a.

In the interval of reactor magnetic flux variation, $\theta_{sa\alpha} - \pi/3 < \theta < \theta_{sa\alpha}$ (Fig. 3), the following equation holds

$$u_{a0} - i_a R = -W_L \omega \frac{d\phi_{a\alpha}}{d\theta}. \tag{25}$$

Simultaneously, equation (2) is valid.

Rectifier $a\alpha$ switches and rectifier $a\beta$ does not switch, since the change in magnetic flux $\phi_{a\beta}$ is connected with the reactors' control circuit. There is a negligible back voltage on rectifier $a\beta$ due to the actual finite impedance of the control winding loop.

$$W_L \omega \frac{d\phi_{a\beta}}{d\theta} < W_L \omega \frac{d\phi_{a\alpha}}{d\theta}.$$

In the saturation interval, $\theta_{sa\alpha} < \theta < \theta_{sa\alpha} + 2\pi/3$, rectifier $a\alpha$ continues to be switched but, on rectifier $a\beta$, there appears the back voltage

$$u_{\mathrm{r}a\beta} = i_a (r + r_{\mathrm{r}}), \tag{26}$$

whose amplitude is easily computed if the magnitude and form of current i_a are known. If there is residual inductance L_2 in the reactors, the magnitude of $u_{ra\beta}$ takes the form:

$$u_{\mathrm{r}a\beta} = i_a \sqrt{(r + r_{\mathrm{r}})^2 + (\omega L_2)^2}. \tag{27}$$

The voltage of rectifier $a\beta$ is shown on Fig. 3,e. On Fig. 8,a there are given the oscillograms of phase current and voltage on one of the rectifiers, which verifies what has been said.

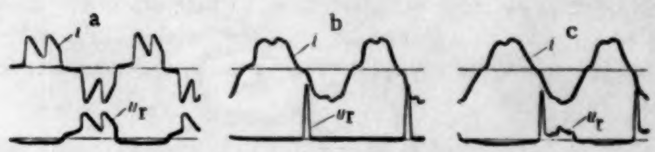


Fig. 8.

In the other modes of magnetic amplifier operation [the modes defined by (4) and (6)], everything so far stated remains in force. Thus, the maximum back voltage on the rectifiers in the case considered occurs with the maximum current in the load circuit.

As an example, we give a comparison of the computed and the measured voltages on the rectifier in the circuit whose parameters were given in Section 1.

The measured circuit impedances were r = 22 ohms, r_r = 2.5 ohms (for a current of 180 milliamperes); the load current for $I_c = 0$ was $I_{L\ max}$ = 180 milliamperes. According to the computations, $U_{r\ max}$ = 4.4 volts while, according to the measurements, $U_{r\ max}$ = 4.5 volts.

Thus, in the given case, the maximum amplitude of the back voltage on the rectifiers was about 0.106 $U_{p\ max}$.

Rectifier operation is different in a magnetic amplifier with suppressed ac components in the control windings. In the interval of increasing magnetic flux $\phi_{a\alpha}$ (Fig. 7), rectifier $a\alpha$ is switched. It remains switched in the interval of reactor $a\alpha$ conductance, and also in the interval $E\theta_k$ where a fall of magnetic flux $\phi_{a\alpha}$ occurs and the following relationship is valid

$$W_{\mathrm{L}} \omega \frac{d\phi_{a\alpha}}{d\theta} > u_{a\alpha'}. \tag{28}$$

Switching of rectifier $a\alpha$ ceases for $\theta = \theta_k$. In the interval $\theta_k < \theta < \theta_{sa\beta}$ the following back voltage is applied to it

$$u_{\mathrm{r}a\alpha} = u_{a\alpha'} - W_{\mathrm{L}} \omega \frac{d\phi_{a\alpha}}{d\theta} \approx u_{a0} - W_{\mathrm{L}} \omega \frac{d\phi_{a\alpha}}{d\theta}. \tag{29}$$

In the interval of reactor $a\beta$ conductance, rectifier $a\alpha$ again begins to be switched, since voltage $u_{aa'}$, for $\theta = \theta_{sa\beta}$, sharply falls from u_{a0} to $u_{a0}(r + r_r)/(R + r + r_r)$ (Fig. 7) and relationship (28) again becomes valid. The oscillograms of the phase current and voltage on the rectifier are given in Fig. 8,b. With small values of control current, the voltage on the rectifier can have a very different character.

In this case, the fall of magnetic flux $\phi_{a\alpha}$, under the action of the control winding's mf, either is concluded in the interval $\theta_k < \theta < \theta_{sa\beta}$ and then, at the end of this interval, $d\phi_{a\alpha}/d\theta = 0$, or proceeds so slowly that the following inequality is valid at the end of this interval

$$u_{a0} \frac{r + r_r}{R_L + r + r_r} > W_L \omega \frac{d\phi_{a\alpha}}{d\theta}. \tag{30}$$

Then, in the interval of reactor $a\beta$ conductance ($\theta_{sa\beta}$ L on Fig. 7), a back voltage equal to

$$u_{ra\alpha} = u_{a0} \frac{r + r_r}{R + r + r_r} - W_L \omega \frac{d\phi_{a\alpha}}{d\theta} \tag{31}$$

remains applied to rectifier $a\alpha$ (Cf., the oscillogram of Fig. 8,c).

As the ratio $(r + r_r)/(R + r + r_r)$ increases, such a character of voltage on the rectifiers remains for larger magnitudes of control current. It should be mentioned that, in an amplifier with free ac components in the control windings, the magnitude of the maximum amplitude of back voltage on the rectifiers is approximately proportional to the ratio $(r + r_r)/(R + r + r_r)$ while, in the given case, its magnitude is virtually independent of this ratio (when the condition $R \gg r + r_r$ holds). This becomes understandable if one takes into account that here the quantity $u_{r\,max}$ is determined in accordance with (29), where none of the quantities in the right member depends on the ratio $(r + r_r)/(R + r + r_r)$.

The absence of any analytic expression for the fall of magnetic flux comes under the action of a constant mf does not permit the computation of the magnitude of the back voltage on the rectifiers for various modes of amplifier operation. Experiments have shown that, in the mode of operation defined by (6), the back voltage on the rectifiers in amplifiers with suppressed ac components in the control windings attains a maximum for saturation angles of the order of $\pi/6$. The amplitude of the back voltage on the rectifiers can reach a value of about $0.7U_{p\,max}$. Thus, one can conclude that the back voltage on the rectifiers in amplifiers with suppressed ac components in the control windings significantly exceeds that in amplifiers with free ac components in the control windings.

By comparing the properties, resulting from our static mode analysis, of the two circuits considered of magnetic amplifiers with self-saturation working on three-phase loads, we can conclude that the circuit with free ac components in the control windings possesses a larger maximum gain and requires a smaller number of rectifying elements, due to the lower back voltage on the rectifiers. The circuit with suppressed ac components in the control windings possesses a better linearity of control characteristic. With the proper choice of the magnitude of the supply voltage, this characteristic can be considered to be virtually linear.

Appendix I

For deriving the functional relationship between control current and saturation angle, we use an idea suggested in [3] for a single-phase amplifier.

We consider a pair of reactors for any of the phases (for example, phase a). If, in some half-period of supply voltage, the magnetic flux of reactor $a\alpha$ increases while the magnetic flux of reactor $a\beta$ decreases, the instantaneous value of the magnetizing component of reactor $a\beta$'s load current is determined from the expression

$$i_{\mu a\beta} = \frac{H_c\, l_c}{W_L}\left(1 - \nu \frac{\phi_{a\beta}}{\Phi_s}\right). \tag{32}$$

By taking (13) into account, we get

$$i_c = -\frac{H_c\, l_c}{W_c}\left(1 - \nu \frac{\phi_{a\beta}}{\Phi_s}\right). \tag{33}$$

The mean value of control current is expressed by the formula

$$i_{\mathrm{C}} = \frac{1}{\pi} \int_{\theta_{\mathrm{I}}}^{\theta_s} i_{\mathrm{C}} d = -\frac{H_{\mathrm{c}}\, l_{\mathrm{c}}}{\pi W_{\mathrm{C}}} \left[(\theta_s - \theta_{\mathrm{I}}) - \frac{\nu}{\Phi_s} \int_{\theta_{\mathrm{I}}}^{\theta_s} \varphi_{a\beta} d\theta \right]. \tag{34}$$

Here, θ_{I} is the angle at which the interval of reactor magnetic flux variation begins.

We consider initially the mode of reactor operation defined by (5). For this mode, the magnetic flux $\Phi_{a\beta}$ is governed by expression (12). By substituting (12) in (34), and taking into account that $u_a = U_m \sin\theta$ and that, in the given case, $\theta_{\mathrm{I}} = \theta_s - \pi/3$, we obtain

$$I_{\mathrm{C}} = -\frac{H_{\mathrm{c}}\, l_{\mathrm{c}}}{\pi W_{\mathrm{C}}} \left[\frac{\pi}{3} - \frac{\nu}{\Phi_s} \int_{\theta_s - \frac{\pi}{3}}^{\theta_s} \left(\Phi - \frac{10^8}{W_{\mathrm{L}}\, \omega} \int_{\theta_s - \frac{\pi}{3}}^{\theta} 1.5 U_m \sin\theta\, d\theta \right) d\theta \right]. \tag{35}$$

After integration, and replacement of U_m by a function of Φ_s in accordance with (1), expression (35) takes the form:

$$I_{\mathrm{C}} = -\frac{H_{\mathrm{c}}\, l_{\mathrm{c}}}{\pi W_{\mathrm{C}}} \left[\frac{\pi}{3} - \nu \left(\frac{\pi}{3} + 0.6 \cos\theta_s - 0.7 \sin\theta_s \right) \right]. \tag{36}$$

For the mode of operation defined by (6), magnetic flux $\Phi_{a\beta}$ has an expression analogous to (12) with the sole difference that the lower limit of integration will be zero ($\theta_{\mathrm{I}} = 0$). Then, by substituting the expression for $\Phi_{a\beta}$ in (34), we obtain

$$I_{\mathrm{C}} = -\frac{H_{\mathrm{c}}\, l_{\mathrm{c}}}{\pi W_{\mathrm{C}}} \left[\theta_s - \frac{\nu}{\Phi_s} \int_0^{\theta_s} \left(\Phi_s - \frac{10^8}{W_{\mathrm{L}}\, \omega} \int_0^{\theta} 1.5\, U_m \sin\theta\, d\theta \right) d\theta \right]. \tag{37}$$

By proceeding with (37) analogously to what we did in the previous case, we obtain, for the mode of reactor operation defined by (6),

$$I_{\mathrm{C}} = -\frac{H_{\mathrm{c}}\, l_{\mathrm{c}}}{\pi W_{\mathrm{C}}} \left[\theta_s - \nu (1.73 \sin\theta_s - 0.73\, \theta_s) \right]. \tag{38}$$

For the mode of operation defined by (4), the reactors' magnetic flux varies, not in accordance with one law, as in the previous cases, but in accordance with three different laws, as a function of the value of the parameter (cf. Equations (20), (25) and (29) in [1]). Therefore, in the right member of (34), instead of

$\int_{\theta_{\mathrm{I}}}^{\theta_s} \varphi_{a\beta}\, d\theta$ we can write the sum of the increments of magnetic flux

$$\sum \Delta\Phi_{a\beta} = \int_{\pi/6}^{\theta_s - \frac{\pi}{3}} \varphi_{a\beta} d\theta + \int_{\theta_s - \frac{\pi}{3}}^{\pi/2} \varphi_{a\beta} d\theta + \int_{\pi/2}^{\theta_s} \varphi_{a\beta} d\theta.$$

The limits of integration correspond to the limits in which equations (20), (25) and (29) of [1] are valid. By substituting in (34) the total increment of magnetic flux, $\Sigma\Delta\Phi_{a\beta}$, which is computed in an analogous fashion to the previous cases by using relationships (2), (5), (7), (20), (25) and (29) from [1] (the explicit expression for $\Sigma\Delta\Phi_{a\beta}$ is not given here for lack of space), we obtain the following expression for the control current

$$I_{\mathrm{c}} = -\frac{H_c l_c}{\pi W_{\mathrm{C}}}\left[\left(\theta_s - \frac{\pi}{6}\right) - \nu(-1.05 + 1.38\sin\theta_s - 0.4\cos\theta_s)\right]. \quad (39)$$

Appendix II

As an example, we give the derivation of the expression for the load current in the mode of reactor operation defined by (5) (Fig. 3,a). In the interval $\theta_{sa\alpha} < \theta < \theta_{sa\alpha} + \pi/3$, the current of phase a is defined by the formula

$$i_a = -\frac{u_{ab}}{2R}, \quad (40)$$

and, in the interval $\theta_{sa\alpha} + \pi/3 < \theta < \theta_{sa\alpha} + 2\pi/3$, by the formula

$$i_a = \frac{u_{ca}}{2R}. \quad (41)$$

The mean value of load current is found from the expression

$$I_{\mathrm{L}} = \frac{1}{2\pi R}\left(-\int\limits_{\theta_s}^{\theta_s + \frac{\pi}{3}} u_{ab}\,d\theta + \int\limits_{\theta_s + \frac{\pi}{3}}^{\theta_s + \frac{2\pi}{3}} u_{ca}\,d\theta\right) =$$

$$= \frac{1}{2\pi R}\left[-\int\limits_{\theta_s}^{\theta_s + \frac{\pi}{3}} \sqrt{3}U_m \sin\left(\theta - \frac{5\pi}{6}\right)d\theta + \int\limits_{\theta_s + \frac{\pi}{3}}^{\theta_s + \frac{2\pi}{3}} \sqrt{3}U_m \sin\left(\theta - \frac{\pi}{6}\right)d\theta\right]. \quad (42)$$

By carrying out the indicated computations, we get

$$I_{\mathrm{L}} = \frac{\sqrt{3}\,U_m}{\pi R}\sin\left(\theta_s + \frac{\pi}{3}\right).$$

The values of load current for the modes defined by (4) and (6) are computed analogously.

Appendix III

We have previously shown the identity of the processes in the load circuits of amplifiers with suppressed ac components in the control windings and the processes in amplifiers with free ac components in the control windings. Therefore, the equations obtained in [1] can also be used in the given case with the absence, referred to above, of the factor 2 in the denominator of the expression for $\Delta\Phi$. We now determine the increment of magnetic flux of reactor $a\alpha$. For the mode of operation defined by (4), this increment is determined as the sum of three increments:

$$\Delta\Phi_{a\alpha} = \Delta\Phi'_{a\alpha} + \Delta\Phi''_{a\alpha} + \Delta\Phi'''_{a\alpha}, \quad (43)$$

since the growth of magnetic flux in the interval $\pi/6 < \theta < \theta_s$ occurs in accordance with three different laws. For $\pi/6 < \theta < \theta_s - \pi/3$, expression (20) is valid, for $\theta_s - \pi/3 < \theta < \pi/2$, expression (25) is valid and, for $\pi/2 < \theta < \theta_s$, expression (29) is valid, all three of these expressions being from work [1].

By substituting the values of the corresponding voltages in these expressions, we obtain

$$\Delta\Phi'_{a\alpha} = \frac{\sqrt{3}\,U_m\cdot 10^8}{W_L\,\omega}\int\limits_{\frac{\pi}{6}}^{\theta_s-\frac{\pi}{3}}\sin\left(\theta-\frac{\pi}{6}\right)\,d\theta = \frac{\sqrt{3}U_m\cdot 10^8}{W_L\,\omega}\,(1-\sin\theta_s), \tag{44}$$

$$\Delta\Phi''_{a\alpha} = \frac{1.5U_m\cdot 10^8}{W_L\,\omega}\int\limits_{\theta_s-\frac{\pi}{3}}^{\pi/2}\sin\theta\,d\theta = \frac{1.5\,U_m\cdot 10^8}{W_L\,\omega}\cos\left(\theta_s-\frac{\pi}{3}\right), \tag{45}$$

$$\Delta\Phi'''_{a\alpha} = \frac{\sqrt{3}\,U_m\cdot 10^8}{W_L\omega}\int\limits_{\frac{\pi}{2}}^{\theta_s}\sin\left(\theta-\frac{5\pi}{6}\right)d\theta = -\frac{\sqrt{3}\,U_m\cdot 10^8}{W_L\omega}\left[\frac{1}{2}+\cos\left(\theta_s+\frac{\pi}{6}\right)\right]. \tag{46}$$

If we substitute (44), (45) and (46) in (43), we obtain, after simplifying the expression

$$\Delta\Phi_{a\alpha} = \frac{\sqrt{3}\,U_m\cdot 10^8}{2W_L\,\omega}\left[1-\cos\left(\theta_s+\frac{\pi}{6}\right)\right]. \tag{22'}$$

For the mode defined by (5), equation (11) is valid, if one takes into account that the upper limit will be θ_s:

$$\Delta\Phi_{a\alpha} = \frac{1.5\,U_m\cdot 10^8}{W_L\omega}\int\limits_{\theta_s-\frac{\pi}{3}}^{\theta_s}\sin\theta\,d\theta = -\frac{1.5\,U_m\cdot 10^8}{W_L\omega}\cos\left(\theta_s+\frac{\pi}{3}\right). \tag{23'}$$

For the mode defined by (6), equation (41) of [1] is valid:

$$\Delta\Phi_{a\alpha} = \frac{1.5\,U_m\cdot 10^8}{W_L\,\omega}\int\limits_{0}^{\theta_s}\sin\theta\,d\theta = \frac{1.5\,U_m\cdot 10^8}{W_L\,\omega}\,(1-\cos\theta_s). \tag{24'}$$

Received August 21, 1958

LITERATURE CITED

[1] A. L. Pisarev, "On the control characteristics of magnetic amplifiers operating on three-phase loads," in Russian], Automation and Remote Control (USSR) 20, 5 (1959).*

[2] H. F. Storm, Magnetic Amplifiers [in Russian translation to English work], ILI (1957).

[3] R. A. Phillips, discussion concerning the paper, H. F. Storm, Theory of Magnetic Amplifiers with Square Loop Core Materials, Trans. AIEE, vol. 72, pt. I, 1953.

[4] L. A. Finzi and D. L. Critchlow, Dynamic Core Behavior of Magnetic Amplifier Performance, Trans. AIEE, vol. 76, pt. I, 1957.

*** See English translation.**

THE USE OF STATISTICAL METHODS FOR DETERMINING THE CHARACTERISTICS OF OBJECTS

SURVEY

Yu. P. Leonov and L. N. Lipatov

(Moscow)

The various statistical methods for determining the characteristics of linear objects are considered.

Recently, in connection with the necessity of increasing control accuracy, the problem of the accurate determination of objects' differential equations has arisen. In particular, several authors believe [1] that lack of ability to determine the characteristics of objects with sufficient exactitude is one of the basic causes limiting the accuracy of control. The best-known methods of determining objects' characteristics, involving the application to their inputs of signals of given forms (sinusoids, step functions, etc.) have essential limitations. On the one hand, they entail the necessity of varying the system's mode of operation. On the other hand, there is ordinarily "noise" in objects (particularly in industrial objects) which either essentially limits the accuracy of measurement or makes it completely impossible. Most recently there have been developed statistical methods for determining the characteristics of objects, these methods being, to a significant degree, free of the disadvantages enumerated above. In this paper we consider various well-known methods of determining object characteristics by statistical means and evaluate them on a comparison basis. In section I we analyze the application of statistical methods for determining frequency characteristics. Although the use of statistical methods ordinarily requires a large volume of computation, they allow the frequency characteristics of systems to be computed in cases where the ordinary methods turn out to be completely inapplicable. In section II we consider methods of determining the equations of objects without disturbing their operating conditions, and we do this for systems with one input, with many inputs, with feedback and for systems which are nonlinear.

I. The Use of Statistical Methods for Obtaining Objects' Frequency Characteristics [3]

1. Preliminary Remarks

The ordinary method for determining amplitude-phase characteristics consists of the following. To the object whose stability is to be investigated a harmonic stimulus ($D \sin \omega t$ or $D \cos \omega t$) is applied; then, if the system is strictly linear, after the natural oscillations of the system have been damped down, there remains an output signal of the form $B \sin (\omega t - \theta)$. The input and output signals are written on one and the same diagram (Fig. 1). One then computes the relative amplitude, $A = B/D$, and the phase shift, θ, i.e., the modulus and phase, respectively, of the investigated object's frequency characteristic for the given frequency ω.

This widely-known method has one essential disadvantage. If the object to be investigated contains an internal noise source, or if additional noise is applied to the object's input, measurement of the output signal becomes either completely impossible or can be carried out only for low frequencies. In this case, the output signal has the form shown in Fig. 2. The noise is particularly powerful in reducing measurement accuracy at high frequencies, when the amplitude of the output signal becomes comparable with the noise amplitude. The

frequency range for which satisfactory results are obtained can be widened by increasing the amplitude of the input signal.

However, all actual objects possess "saturation" and the last measure can, therefore, only partially ameliorate the situation. It has been established in practice that the highest frequency for which the amplitude of the output signal of a "noisy" object can be measured is obtained when the amplitude of the output signal equals the mean square value of the noise.

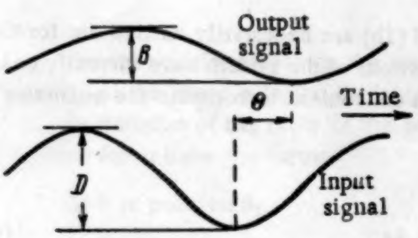


Fig. 1.

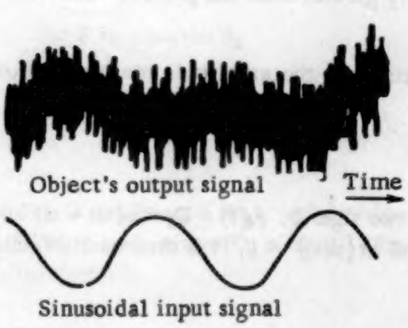


Fig. 2.

To overcome the difficulties stated above, we consider statistical methods of determining frequency characteristics. One of them (the amplitude method) makes it possible to determine, when there is noise in the object, the imaginary and real parts of the frequency characteristic. The second (the method of null phase) is based on phase measurements. These methods require the mechanization of the computational work. However, they are applicable in cases where the ordinary methods are completely inapplicable.

2. The Amplitude Method

This method is based on the following relationships. Let there be two signals, $f_3(t) = D \sin \omega t$ and $f_2(t) = B \sin(\omega t + \theta) + n(t)$, where $n(t)$ is a stationary random function (the object's "noise") and $M\{n(t)\} = 0$, M being the symbol for mathematical expectation.

It is then easily verified that

$$R_{3,2}(0) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f_3(t) f_2(t)\, dt = +\frac{1}{2} BD \cos\theta, \tag{1}$$

Correspondingly,

$$R_{1,2}(0) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f_1(t) f_2(t)\, dt = \frac{1}{2} BD \sin\theta, \tag{2}$$

where $f_1(t) = D \cos \omega t$.

On the basis of (1) and (2) we have

$$A = \frac{2}{D^2} \sqrt{R_{1,2}^2(0) + R_{3,2}^2(0)}, \quad \theta = \operatorname{arc\,tan} \frac{R_{1,2}(0)}{R_{3,2}(0)}, \tag{3}$$

where $A = B/D$ is the value of the amplitude characteristic for frequency ω and θ is the value of the phase characteristic at the same frequency.

We note that $R_{1,2}(0)$ and $R_{3,2}(0)$ are the values of the cross-correlation functions $R_{1,2}(\tau)$ and $R_{3,2}(\tau)$ for $\tau = 0$. It is easily seen, moreover, that $R_{1,2}(0)$ and $R_{3,2}(0)$ are proportional, respectively, to the real and imaginary parts of the system's frequency characteristic vector at the frequency ω.

It is completely obvious that computing $R_{1,2}(0)$ and $R_{3,2}(0)$ exactly is impossible because of the impossibility of implementing the passage to the limit as $T \to \infty$. In actuality, computation is not done by formulae (1) and (2) but by the formulae

$$\overline{R}_{3,2}(0) = \frac{1}{T} \int_0^T f_3(t) f_2(t)\, dt, \tag{1a}$$

$$\overline{R}_{1,2}(0) = \frac{1}{T}\int_0^T f_1(t)\, f_2(t)\, dt. \tag{1b}$$

The quantities $\overline{R}_{1,2}(0)$ and $\overline{R}_{3,2}(0)$ are estimates of $R_{1,2}(0)$ and $R_{3,2}(0)$. These estimates are the more accurate, the longer the time of integration T. In particular, it is advantageous to take $T = k\, 2\pi/\omega$, where k is a sufficiently large positive integer. We note that $\overline{R}_{1,2}$ and $\overline{R}_{3,2}$ depend on T. However, if the interval of integration is chosen sufficiently large, $\overline{R}_{1,2}$ and $\overline{R}_{3,2}$ do not, for all practical purposes, depend on T. This serves as a criterion for the correctness of the choice of T.

It is essential to remark that the computations of integrals (1a) and (1b) are necessarily carried out for the function $f_2(t)$ for $t > t_1$, where t_1 is a time after which the natural oscillations of the system have virtually ceased. Since these computations can be carried out on an analog computer, it is convenient to compute the estimates of $\overline{R}_{1,2}(0)$ and $\overline{R}_{3,2}(0)$ from the formula

$$\overline{R}_{1,2}(0) \approx \frac{1}{T_0}\int_0^t e^{\frac{-(t-x)}{T_0}} f_1(x)\, f_2(x)\, dx, \tag{4}$$

where T_0 is the computer's (model's) time constant. If T_0 is significantly greater than the period, $2\pi/\omega$ of the input signal, then $\overline{R}_{1,2}(0)$ is virtually independent of t for $t > 5T_0$.

The basic scheme necessary for obtaining the frequency characteristic by the amplitude method is shown in Fig. 3.

3. The Method of Null Phase

This method is based on the following relationships. If there are two signals, $f_4(t) = D_1\cos(\omega t + \alpha)$ and $f_2(t) = B\sin(\omega t + \theta) + n(t)$, where n(t) is a stationary random function and $M\{n(t)\} = 0$, it is then easily found that

$$R_{4,2}(0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T f_4(t)\, f_2(t)\, dt = \frac{1}{2} D_1 B \sin(\theta - \alpha). \tag{5}$$

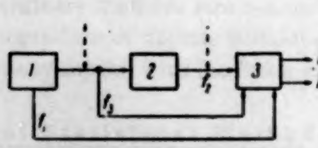


Fig. 3. 1) Sine and cosine generator, 2) object and 3) computer.

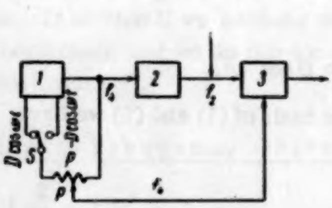


Fig. 4. 1) Sine and cosine generator, 2) object and 3) computer.

We now assume, as we earlier did, that $f_2(t)$ is the output signal of an object at whose input the signal $f_3(t) = D\sin\omega t$ has been applied and at whose output there occurs the noise n(t). Then, if the phase, α, of signal $f_4(t)$ is made equal to θ then, in accordance with (5), we obtain

$$R_{4,2}(0) = 0 \tag{6}$$

for

$$\alpha = \theta. \tag{7}$$

Thus, by knowing the established phase α for which $R_{4,2}(0) = 0$, we obtain the phase of the frequency characteristic for frequency ω. In order to obtain the modulus of the frequency characteristic, we initially establish a value of α for which $R_{4,2}(0) = 0$, and then we so vary it that the algebraic sum of the phases becomes equal to 90°.

Then on the basis of (5), we have $R_{4,2}(0) = D_1B/2$, from whence

$$A = \frac{B}{D} = 2\frac{R_{4,2}(0)}{DD_1}. \tag{8}$$

The scheme for obtaining the frequency characteristic by the null phase method is shown in Fig. 4. A smooth variation of the phase α in the limits $(0, \pm\pi/2)$ is attained by means of potentiometer P and switch S. The value of α as a function of the ratio of the potentiometer arms, P_b, can be obtained from the diagram of Fig. 5. The formulae for α have the forms:

for S in position S_1

$$\alpha = -\operatorname{arc\,tan}\left(\frac{1-P_b}{P_b}\right); \tag{9}$$

for S in position S_2

$$\alpha = -\frac{\pi}{2} + \operatorname{arc\,tan}\left(\frac{1-P_b}{P_b}\right).$$

It is easily seen that, as P_b varies, there is a corresponding change in the amplitude of the signal $f_4(t) = D_1\cos(\omega t + \alpha)$, where $D_1 = m_bD$. The correction factor m_b for taking account of the variations in amplitude D_1 has the form:

$$m_b = \sqrt{1 - 2P_b + 2P_b^2} \tag{10}$$

The curve for m_b is shown in Fig. 6.

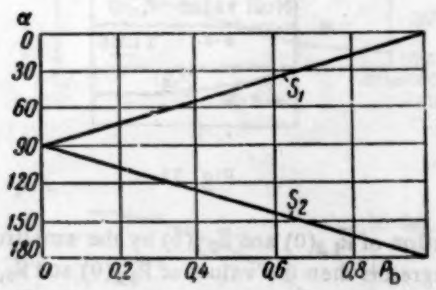


Fig. 5.

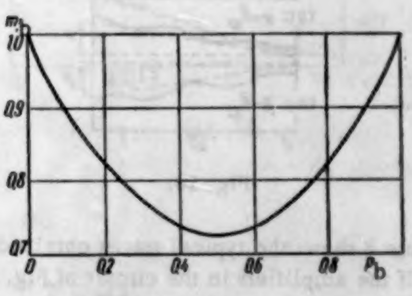


Fig. 6.

Because of the variations in amplitude D_1, it is convenient, in computing the amplitude characteristic, to use the formula

$$A = 2\frac{R_{4,2}(0)}{D^2 m_b}.$$

4. A Practical Circuit for Obtaining Frequency Characteristics by Statistical Methods

The quantities $\overline{R}_{1,2}(0)$, $\overline{R}_{3,2}(0)$ and $\overline{R}_{4,2}(0)$ can be easily computed by means of an electrical model (analog) for solving differential equations. The functional schematic of the analog for determining the magnitude of these quantities is shown in Fig. 7. The feedback paths around the amplifiers allow the time constants to be varied within wide limits. Ordinarily, one chooses $T_1 = T_2$.

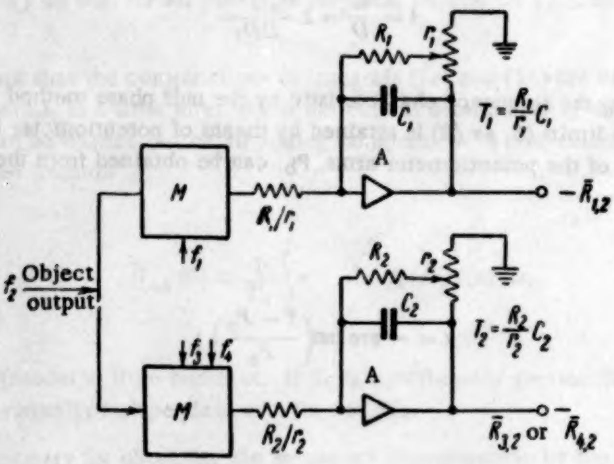


Fig. 7. M is a multiplier, A is an operational amplifier, and r is the potentiometer ratio.

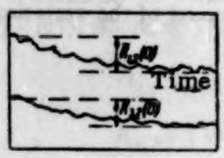


Fig. 8.

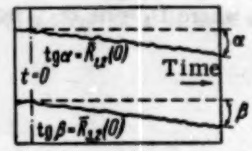


Fig. 9.

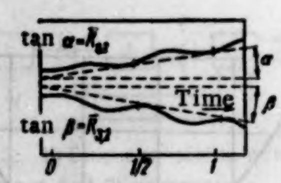


Fig. 10.

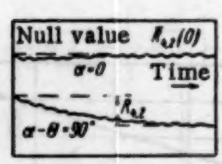


Fig. 11.

Figure 8 shows the typical traces obtained in the determination of $\overline{R}_{1,2}(0)$ and $\overline{R}_{3,2}(0)$ by the amplitude method. If the amplifiers in the circuit of Fig. 7 are used as integrators then the values of $\overline{R}_{1,2}(0)$ and $\overline{R}_{3,2}(0)$ equal the mean slopes of the line obtained in this case. Characteristic traces are shown in Fig. 9. For very low frequencies of the sinusoidal input signal, a trace such as shown in Fig. 10 is obtained. With this, it is necessary to produce the average lines as is shown on the diagram. A typical trace of the determination of the magnitude of $\overline{R}_{4,2}(0)$ by the null phase method is shown on Fig. 11.

For the methods considered, it is necessary that the sinusoidal and cosinusoidal signals have one and the same frequency and amplitude. These signals can be obtained either by means of a generator for sin ωt or by means of a circuit set up on the model of Fig. 12. It is certainly convenient to obtain these functions by means of a model (analog). This can be done by solving the equation

$$\frac{d^2x}{dt^2} + \omega^2 x = 0. \tag{11}$$

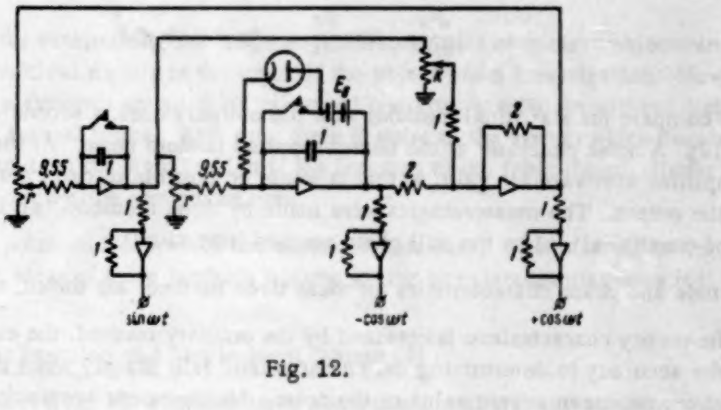


Fig. 12.

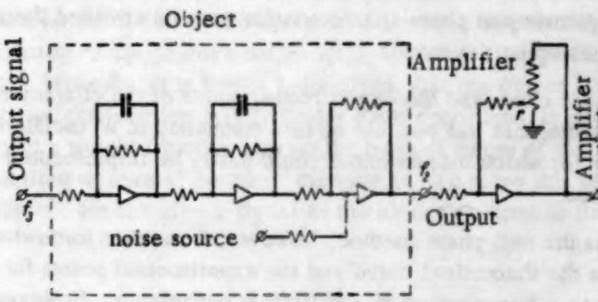


Fig. 13.

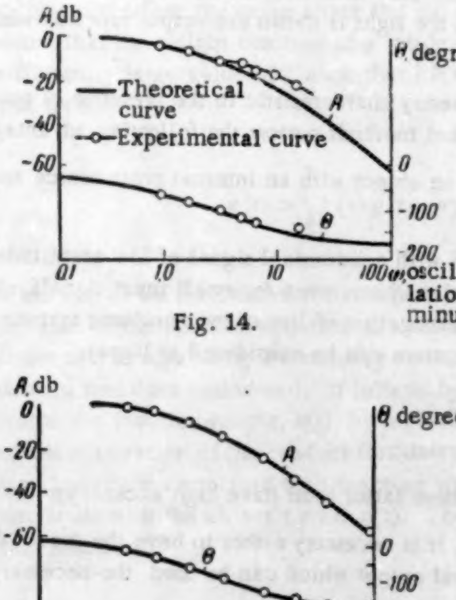


Fig. 14.

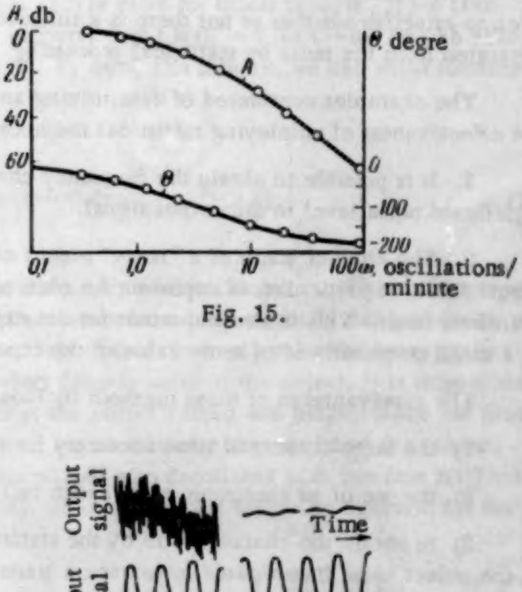


Fig. 15.

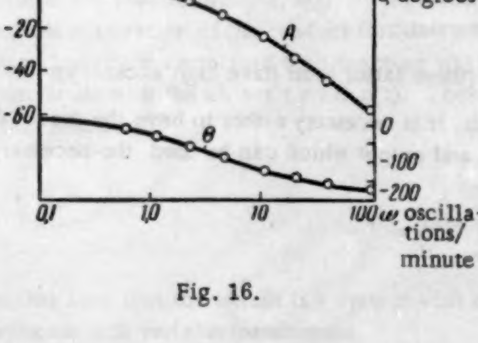


Fig. 16.

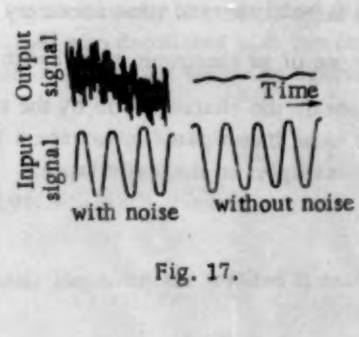


Fig. 17.

on the model. In practice, in order to obtain generation of a model, it is necessary to set up an equation with a negative damping coefficient, of the form

$$\frac{d^2x}{dt^2} - 2k\omega\frac{dx}{dt} + \omega^2 x = 0, \ k \ll 1. \tag{12}$$

In order to compare the statistical methods with the ordinary ones, a second-order equation was set up on the model (Fig. 13). A noise generator at the object supplied random noise. At the object's output there was connected an amplifier with variable gain, so that it would be possible to carry out measurements for small values of amplitude at the output. The measurements were made by three methods: a) by the ordinary methods, b) by the amplitude method (statistical), c) by the null phase method (statistical).

The amplitude and phase characteristics for these three methods are shown, respectively, in Figs. 14-16.

When the frequency characteristic is obtained by the ordinary method, the conclusion can be drawn, based on Fig. 14, that the accuracy in determining the characteristic falls sharply when the output amplitude is made less than, or equal to, the mean square value of the noise. Measurement accuracy is good only up to a frequency of 4 oscillations a minute (0.07 cycles), for which the amplitude of the output signal equals the mean square value of the noise. The exact amplitude and phase characteristics were determined theoretically on the basis of a well-known equation, which was set up on the model.

Figure 15 gives the exact curve and the experimental points of the characteristic as obtained by the amplitude method. Accurate measurement was possible up to a frequency of 95 oscillations per minute. Thus, the upper limit on the frequency for which measurement could easily be implemented was 20 times higher than this limit with the ordinary method.

The most accurate was the null phase method, however, it requires somewhat more time for the measurements to be made. Figure 16 gives the theoretical curve and the experimental points for this case. Accurate measurement here was also carried out up to a frequency of 95 oscillations per minute. However, the accuracy of the null phase method is, in practice, higher than that of the amplitude method, since the null phase method is a compensating one. The superiority of the statistical method is particularly clearly seen on Fig. 17, where the input and output signals are given with the presence of noise, and without it. It is clear from Fig. 17 that the sinusoidal signal at the output is completely hidden in the noise. It is impossible, not only to determine its amplitude and phase, but even to establish whether or not there is a sinusoidal signal. On the right is shown the output function which was separated from the noise by statistical processing.

The examples considered of determining an object's frequency characteristic in the presence of noise proves the effectiveness of employing statistical methods. The statistical methods possess the following advantages.

1. It is possible to obtain the frequency characteristic of an object with an internal noise source and with a significant noise level in the output signal.

2. The characteristic of a "noisy" object can be obtained with a sinusoidal signal of low amplitude at the input. This, in particular, is important for objects with large gains where, even for small input signals, saturation manifests itself. This is also important for the experimental investigation of linearized nonlinear systems where, in a small neighborhood of some value of the input signal, the system can be considered as linear.

The disadvantages of these methods include the following:

1) the large amount of time necessary for taking measurements;

2) the use of an electronic model with two multipliers (these latter must have high accuracy);

3) to obtain the characteristic by the statistical methods, it is necessary either to have the model directly at the object to be investigated or to have a trace of its input and output which can be used the necessary number of times (for example, on magnetic tape).

II. Determination of Objects' Characteristics While They are in Operation [4-9]

1. Preliminary Remarks

The statistical methods for determining the frequency characteristics of "noisy" objects are very convenient when it is possible to apply sinusoidal signals to the input of the object under investigation. However, it is frequently required to determine a system's equation in its normal operating conditions without disturbing any connections or applying signals of special forms. With this, there is noise in the system which does not permit the registering, at the system's output, of the signal which is the reaction to the input signal. Under these conditions, the methods given in the previous section can not be used.

We will consider at this point methods which are useful for this case. As was stated in the introduction, we shall present formally only the ideas of these methods without giving any rigorous mathematical foundation for them.

2. Determining the Weight Function of a Single Input System [4]

We consider a linear object (Fig. 18) at whose input there is a signal y(t) and at whose output there is a signal x(t). In the object itself there is a noise source which engenders an equivalent noise signal at the output. It is assumed that the statistical characteristics of the noise are unknown. In particular, it is impossible to obtain a trace (registration) of a realization of the object's noise n(t). Moreover, it is not known what effect the object's noise n(t) has on the signal x(t). However, it is known beforehand that the object is linear and, consequently, is completely characterized by its weight function k(t).* Under these conditions, it is necessary to determine the object's weight function k(t) on the basis of traces of the realization of random functions at several points of the system. To solve this problem, we can proceed as follows. We imagine a signal at the object's output in the form

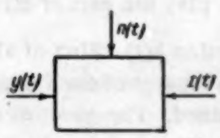


Fig. 18.

$$x(t) = \int_0^t k(\tau)\, y(t-\tau)\, d\tau + \int_0^t k_n(\tau)\, n(t-\tau)\, d\tau, \tag{13}$$

where k(t) is the system's weight function between input and output and $k_n(t)$ is the system's weight function between the point where the noise arises and the output. Formula (13) is valid for linear objects. If we take into account that the weight function of a stable object has the property that $|k(t)| \to 0$ for $t \to \infty$, we can always find a sufficiently large value of T such that $|k(t)| \approx 0$ for $t \geq T$. By using this remark, we can write formula (13) in the form

$$x(t) \approx \int_0^T k(\tau)\, y(t-\tau)\, d\tau + \int_0^T k_n(\tau)\, n(t-\tau)\, d\tau, \qquad (t \geqslant T). \tag{14}$$

In the sequel we shall assume that signals y(t) and n(t) are stationary random functions. Then, if we consider the value of the function x(t) after the natural motions of the system have been virtually damped down, we can consider x(t) as also being a stationary random function. Formula (14) is obviously valid for every realization of the random functions considered. It follows from this that, when there is noise in the object, it is impossible to determine the function sought, k(t), by registering the signals at the object's input and output, since the function $k_n(t)$ and the realization of the random function n(t) are unknown. We now assume that we know the random function z(t) which is correlated with function y(t) [and, consequently, is also correlated with function x(t)] but is not correlated with the object's noise n(t). Then, by using (14), we can write the following equation for the correlation functions:

$$R_{xz}(t) = \int_0^T k(\tau)\, R_{yz}(t-\tau)\, d\tau, \tag{15}$$

* We assume here that the object is a system with constant parameters. However, the method is easily applied to linear systems with variable parameters.

where $R_{xz}(t) = \lim_{T_1 \to \infty} \frac{1}{T_1} \int_0^{T_1} x(u)\, z(u-t)\, du$ is the cross-correlation function of x(t) and z(t) and $R_{yz}(t)$ is the cross-correlation function of y(t) and z(t). With the conditions assumed, cross-correlation function $R_{nz}(t) = 0$ and the second integral in equation (14) reduces to zero. In particular, in many problems it can be assumed that n(t) is not correlated with the input function y(t). Then, we can set z = y and rewrite (15) in the form

$$R_{xy}(t) = \int_0^T k(\tau)\, R_{yy}(t-\tau)\, d\tau, \tag{15a}$$

where $R_{yy}(t) = \lim_{T_1 \to \infty} \frac{1}{T_1} \int_0^{T_1} y(u)\, y(u-t)\, du$ is the auto-correlation function of process y(t). However, it is necessary to bear in mind that y(t) is not the only random function which permits relationship (15) to be obtained.

Equation (15) gives the solution of the problem posed. By solving it, we can determine k(t), the function sought.

Thus, for the solution in the general case it is necessary to have the traces of the realizations of three functions, z(t), y(t) and x(t) but, if it is assumed that y(t) and n(t) are uncorrelated, then only the traces of two functions are necessary since one can set y(t) = z(t). We note that the function x(t) cannot, obviously, play the part of z(t).

In conclusion we can make several remarks as to the computation of weight function k(t). First of all, it is impossible to compute $R_{yy}(t)$ and $R_{xy}(t)$ exactly in substituting them in equation (15a) because of the finite duration of the traces. Only estimates, $\overline{R}_{yy}(t)$ and $\overline{R}_{xy}(t)$, of these functions can be obtained. The question of the accuracy of determining k(t) because of errors in determining the functions $R_{yy}(t)$ and $R_{xy}(t)$ is beyond the limits of this work. It is quite obvious, however, that the more accurately $\overline{R}_{yy}(t)$ and $\overline{R}_{xy}(t)$ are computed, the closer $\overline{k}(t)$ will be to the solution k(t). The errors in $\overline{R}_{yy}(t)$ and $\overline{R}_{xy}(t)$ due to the finite duration of the traces of the realizations of x(t) and y(t) can be easily determined by very well-known methods [2].

3. Determination of the Weight Function of a Multi-Input Linear System [4]

The case frequently occurs when it is necessary to determine the weight function of a system with several inputs and several outputs. The method considered for determination of the weight function is easily extended to this case.

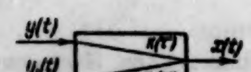


Fig. 19.

A system with several outputs does not give rise to any additional difficulties since each output can be considered independently of the others. Consequently, a system with n outputs can be considered, as n independent systems have the same inputs. Therefore, we can limit our consideration to systems with several inputs and one output. Initially we consider a system with two inputs (Fig. 19). Such a system can be characterized by two weight functions, k(t) and $k_1(t)$. We will assume, as before, that $k(t) = k_1(t) = 0$ for $t \geq T$, where T is the system's memory.* For a two-input system, we can obviously write the following expression for the output signal:

$$x(t) = \int_0^T k(\tau)\, y(t-\tau)\, d\tau + \int_0^T k_1(\tau)\, y_1(t-\tau)\, d\tau. \tag{16}$$

From this we obtain the system of two integral equations:

$$R_{yx}(t) = \int_0^T k(\tau)\, R_{yy}(t-\tau)\, d\tau + \int_0^T k_1(\tau)\, R_{yy_1}(t-\tau)\, d\tau, \tag{17a}$$

* The "memory" is the interval OT in which the weight function k(t) is not identically zero.

$$R_{y_1x}(t)=\int_0^T k(\tau)\,R_{y_1y}(t-\tau)\,d\tau+\int_0^T k_1(\tau)\,R_{y_1y_1}(t-\tau)\,d\tau. \tag{17b}$$

If $R_{y_1y}(t)=0$ then, to determine k(t) and $k_1(t)$ we obtain two separate equations analogous to equation (15). If, however, y and y_1 are correlated, we can then separate from signal $y_1(t)$ a component z(t) which is not correlated with y(t). Indeed, we have the equation

$$y_1(t)=z(t)+\int_{-\infty}^{+\infty} k_3(\tau)\,y(t-\tau)\,d\tau. \tag{18}$$

On the basis of (18) we obtain, for the correlation functions,

$$R_{yy_1}(t)=R_{yz}(t)+\int_{-\infty}^{+\infty} k_3(\tau)\,R_{yy}(t-\tau)\,d\tau. \tag{19}$$

It is clear from this last expression that, if there exists a solution $k_3^*(t)$ of the integral equation

$$R_{yy_1}(t)=\int_{-\infty}^{+\infty} k_3(\tau)\,R_{yy}(t-\tau)\,d\tau, \tag{20}$$

then, by setting $k_3(t)=k_3^*(t)$, we obtain, from (19), $R_{yz}(t)=0$. The presentation of signal $y_1(t)$ in the form of (18), where z(t) is not correlated with y(t) can be used to reduce the solution of the system of integral equations (17a) and (17b) to the solution of three separated equations. In fact, on the basis of (18) we have

$$R_{y_1y_1}(t)=R_{zz}(t)+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k_3(\tau_1)\,k_3(\tau_2)\,R_{yy}(t-\tau_1+\tau_2)\,d\tau_1 d\tau_2= \\ =R_{zz}(t)+\int_{-\infty}^{+\infty} k_3(\tau)\,R_{yy_1}(t+\tau)\,d\tau, \tag{21}$$

$$R_{y_1x}(t)=R_{zx}(t)+\int_{-\infty}^{+\infty} k_3(\tau)\,R_{yx}(t-\tau)d\tau. \tag{22}$$

If we substitute $R_{y_1y}(t)$ from (20) and $R_{y_1y_1}(t)$ from (21) in (17b), we get

$$R_{y_1x}(t)=\int_0^T\int_{-\infty}^{+\infty} k(\tau_1)\,k_3(\tau_2)\,R_{yy}(t+\tau_2-\tau_1)\,d\tau_1 d\tau_2+ \\ +\int_0^T k_1(\tau_1)\,R_{zz}(t-\tau_1)\,d\tau_1+\int_0^T\int_{-\infty}^{+\infty} k_1(\tau_1)\,k_3(\tau_2)\,R_{yy_1}(t+\tau_2-\tau_1)\,d\tau_1 d\tau_2. \tag{23}$$

By substituting (17a) in (22), we obtain

$$R_{y_1x}(t)=R_{zx}(t)+\int_{-\infty}^{+\infty} k_3(\tau_1)\,d\tau_1\int_0^T [k(\tau_2)\,R_{yy}(t+\tau_1-\tau_2)+ \\ +k_1(\tau_2)\,R_{yy_1}(t+\tau_1-\tau_2)]\,d\tau_2. \tag{24}$$

By comparing (23) and (24) we obtain

$$R_{zx}(t) = \int_0^T k_1(\tau_1) R_{zz}(t - \tau_1)\, d\tau_1. \tag{25}$$

If we interchange the roles of y and y_1, we can obtain an analogous equation of k(t). For using (25), functions $R_{zx}(t)$ and $R_{zz}(t)$ are obtained from (21) and (22). We note that equations (17a) and (17b) as well as the method for determining the weight functions k(t) and $k_1(t)$ remain in force if, in the system (Fig. 19), there is a source of noise which is not correlated with y(t) and $y_1(t)$.

4. Determining the Characteristic of a System with Feedback [4, 5]

Let there now be a complex multi-loop system (Fig. 20) consisting of many elements, each of which can contain a source of noise. As before, we assume that it is impossible to obtain directly the exhaustive characteristics of the noise arising in any element of the system. Assuming the elements to be linear, we must determine the system's weight function on the basis of the traces of realizations obtained at definite points of the system. As for an object which is not shunted by a feedback path, the question arises as to which points of the system must be tapped for traces of realizations and how many of such realizations are necessary to compute the weight function. To solve this problem we use the same method as in the case considered in section 2, namely, we register the realizations of three signals: y(t) at the element's input, x(t) and some auxiliary signal z(t) at the element's output. With this, as before, signal z(t) must be correlated with y(t) but not with the noise n(t) arising within the element. If it is then necessary, for example, to determine the weight function of element 3 (Fig. 20), integral equation (15) must be solved for the latter. However, the essential difference between objects shunted by feedback paths and open-loop systems consists in the choice of the function z(t). In particular, for an element shunted by feedback paths it is no longer possible to choose the input function y(t) as the function z(t). Due to the feedback, y(t) will necessarily be correlated with the noise arising in the element investigated and, consequently, it cannot be eliminated by Equation (14). For the scheme of Fig. 20 it is possible, in many cases, to choose function $l_1(t)$ or $l_2(t)$ as z(t). In particular, the method to be presented is applicable to the determination of the weight function of an object in a single-loop system (Fig. 21). As signal z(t) one can frequently use the signal at the input of the system (Fig. 21), but not at the input of the link being investigated.

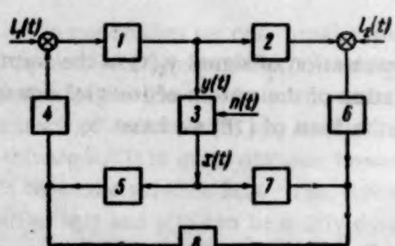


Fig. 20.

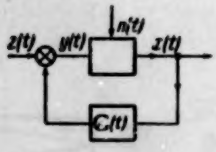


Fig. 21.

We now consider, as applied to the single-loop system (Fig. 21), the case when it is impossible to find a function z(t) which satisfies the necessary requirements. It is then possible to determine the object's weight function by carrying out an experiment twice for different controller parameters.

By denoting the functions of the first experiment by the superscript 1 and those of the second by the superscript 2, we can easily obtain the following formula [6]:

$$\Phi(j\omega) = \frac{{}^2G_{xx}^{+}(j\omega) - {}^1G_{xx}^{+}(j\omega)}{\Phi_{C1}(j\omega)\,{}^1G_{xx}^{+} - \Phi_{C2}(j\omega)\,{}^2G_{xx}^{+}(j\omega)}, \tag{26}$$

where $\Phi(j\omega)$ is the object's frequency characteristic and $\Phi_{C1}(j\omega)$ and $\Phi_{C2}(j\omega)$ are the controller's transfer functions for the first and second tunings respectively. The function ${}^1G_{xx}^{+}(j\omega)$ is defined by the formula

$${}^1G_{xx}(j\omega) = {}^1G_{xx}^{+}(j\omega)\,{}^1G_{xx}^{-}(j\omega).$$

where ${}^1G_{XX}(j\omega)$ is the spectral density function of the output signal for the first tuning of the controller. By using this method we can also easily obtain the spectral density function of the object's noise.

Thus, by knowing the statistical characteristics of the signals x(t) for two trials and by knowing Φ_{C1} and Φ_{C2}, we can determine the object's transfer function and the spectral density of the noise G_{nn} for the case of the system of Fig. 21. However, the method considered possesses two disadvantages. First, intervention in system operation is required, since it is necessary to vary the controller's parameters. Second, the duration of the experiment is significantly increased and, during this time, the object's characteristics may change.

5. Determining the Best Linear Approximations of Nonlinear Systems [4, 5, 9]

Up untill now we have only considered linear systems, which are completely characterized by their weight function k(t). We note that the methods presented in section II can easily be extended to the case of nonstationary input signals. In this case we obtain more complicated integral equations for determining the weight function k(t), which will be a function of the variables t and τ:

$$R_{xy}(t,\ u) = \int_0^T k(t,\ \tau)\ R_{yy}(u,\ t-\tau)\, d\tau,$$

where

$$R_{yy}(t_1,\ t_2) = M\,[Y(t_1)\, Y(t_2)]$$

and

$$R_{xy}(t_1,\ t_2) = M\,[X(t_1)\, Y(t_2)].$$

The method of determining the weight function of linear systems can be applied for a constructed linear system, the output signal of which is close (in the sense of a minimum second error moment) to the signal at the output of the given nonlinear system. Indeed, let the output signal of the nonlinear system be x(t) and the input signal y(t). We can then determine the equivalent linear system for the condition that the second error moment, σ^2, be a minimum, namely,

$$\sigma^2 = M\left\{x(t) - \int_0^T h(\tau)\, y(t-\tau)\, d\tau\right\}^2. \tag{27}$$

By applying the well-known methods of variational calculus to determine the h(t) which minimizes the functional σ^2 after computation of the mathematical expectation, we obtain the integral equation

$$R_{yx}(t) = \int_0^T h(\tau)\, R_{yy}(t-\tau)\, d\tau. \tag{28}$$

It is easily seen that the function h(t) which satisfied (28) minimizes the function σ^2. In certain problems, the approximation just described of nonlinear systems by means of linear systems close to the nonlinear (in the sense cited) is very convenient. In particular, many nonlinear systems admit linearization. This means that, for small disturbances in the neighborhood of some value of the signal, they behave as linear systems. Thus, the quality and stability of such systems can be determined on the basis of a linear model with the condition that the disturbances be small. For this, the random disturbances must be small.

If the random disturbances in the system's operating processes are large, the equivalent linear system, obtained by processing such traces, is a statistically linearized system. This means that all the nonlinearities in the nonlinear system are replaced by equivalent linear elements whose coefficients depend on the mean square value of the noise. With this, for each value of noise dispersion in the input signal, the coefficients are chosen from the condition that the signal at the system's output best approximate, on the average, the output signal of the linear system.

With a sinusoidal input signal, a statistically linearized system coincides with a linear system computed on the basis of a harmonic output signal. In this case, statistical linearization is equivalent to ignoring all the higher harmonics. Despite the convenience of representing a nonlinear system by an equivalent linear system, the latter does not completely reproduce all the properties of the nonlinear system and can therefore not replace it in all cases.

6. Several Remarks on the Solution of the Basic Integral Equation [4, 5, 7, 8]

The solution of the basis integral equation poses no difficulties. One method of solving it consists of transforming equation (15a) to a system of linear algebraic equations by replacing the integral by a sum. We then obtain

$$\sum_{\mu=1}^{n} a_{\nu\mu} k_{\mu} = b_{\nu} \quad (\nu = 1, \ldots, n), \tag{29}$$

where $k_{\mu} = k(\tau_{\mu})$, $b_{\nu} = R_{xy}(t_{\nu})$, $a_{\nu\mu} = R_{yy}(t_{\nu} - \tau_{\mu})$. The system of linear algebraic equations can be solved either on a general-purpose computer or on a special-purpose machine. The method of transforming an integral equation to a system of linear algebraic equations was used in work [4]. A second method of solving equation (15a) amounts to the finding of a continuous solution by the method of successive approximations. This method was investigated as applied to optimal systems in [7, 8]. This method is applicable, without change, to the finding of an approximate solution of equation (15a). To determine the successive approximations, as was shown in [7, 8], one must use the formula

$$k_{n+1}(t) = k_n(t) - \alpha \left[\int_0^T R_{yy}(t-\tau)\, k_n(\tau)\, d\tau - R_{xy}(t) \right] \quad (n = 0, 1, \ldots). \tag{30}$$

The initial (null) approximation, $k_0(t)$, in formula (30) is chosen arbitrarily. The number α must be chosen from the condition that $\alpha \leq 1/\lambda_{max}$ where λ_{max} is the largest eigenvalue of the kernel, the correlation function $R_{yy}(t)$. In practice, one can take $\alpha = 1/2$ in the majority of problems. We note that the algorithm of (30) is a variant of the method of steepest descent. However, the computations connected with (30) are much simpler than those used in the algorithm for the method of steepest descent which has been applied until recently. This simplification results from the fact that, in (30), α is chosen as α = const whereas, in the usual formula for the method of steepest descent, it is necessary to compute α painfully at each step. When formula (30) is used, convergence is somewhat slower than with the formula for steepest descent, but this decrease in speed of convergence is very significant and is completely compensated for by the choice of a constant α.

Finally, for solving equation (15a) formally, the method of imaging (transforms) may be used.

By using the two-sided Laplace transform, we obtain, on the basis of (15a),

$$R_{xy}(p) = K(p)\, R_{yy}(-p^2), \tag{31}$$

where $R_{xy}(p)$, $K(p)$ and $R_{yy}(-p^2)$ are the Laplace transforms of, respectively, the functions $R_{xy}(t)$, $k(t)$ and $R_{yy}(t)$. From (31) we have

$$K(p) = \frac{R_{xy}(p)}{R_{yy}(-p^2)}. \tag{32}$$

It is essential to note that, in (31), K(p) will be the transfer function of a physically realizable system. However, in the author's opinion, formula (32) is impossible to use for direct calculations. Indeed, in practical cases, there are no functions $R_{xy}(p)$ and $R_{yy}(-p^2)$, but estimates* of these functions, $\overline{R}_{xy}(p)$ and $\overline{R}_{yy}(-p^2)$. Moreover, the processes for which the corresponding estimates are computed are only asymptotically stationary. All this leads to the circumstance that, instead of formula (32), one uses the formula

*The estimates $\overline{R}_{xy}(p)$ and $\overline{R}_{yy}(-p^2)$ of the functions $R_{xy}(p)$ and $R_{yy}(-p^2)$ are called random functions close, in some sense or other, to $R_{xy}(p)$ and $R_{yy}(-p^2)$.

$$\overline{K}(p) = \frac{\overline{R}_{xy}(p)}{\overline{R}_{yy}(-p^2)}, \tag{33}$$

where the superscript bar denotes an estimate of the corresponding function. It is obvious that, from (33), one can obtain a $\overline{K}(p)$ which corresponds, in particular, to a physically unrealizable unstable system. In this inheres the difficulty in a direct usage of formula (33). This difficulty can be avoided if, for solving the problem, we use, not formula (33), but a formula from the theory of optimal systems. It is well-known [2] that in optimal systems theory one solves an equation of the same form as in (15). However, the essential difference between the solution of the problem for optimal systems and the solution of equation (15) consists in the fact that, in the first case, the functions $R_{xy}(t)$ and $R_{yy}(t)$ are given, and it is required to determine the weight function which lies in the class of physically realizable systems, while in the second case it is known beforehand that the system is physically realizable and the function X(t) is obtained as the result of passing the random function Y(t) through the system. It is obvious that the method of solving the equation for the optimal system is also useful for finding the weight function which satisfied equation (15a). With this, the convenience of using the results which relate to optimal systems consists in this, that with realizations $\overline{R}_{xy}(t)$ and $\overline{R}_{yy}(t)$ of the correlation functions $R_{xy}(t)$ and $R_{yy}(t)$, a solution is obtained which corresponds to a stable and physically realizable system. The realizations $\overline{R}_{xy}(t)$ and $\overline{R}_{yy}(t)$ are computed by very well-known methods on the basis of realizations of the processes X(t) and Y(t).

SUMMARY

The statistical methods considered in section II have great potentialities as compared with the ordinary methods for finding frequency characteristics. In particular, the most important distinguishing feature of these methods is the possibility they offer of determining the characteristic of a system without disturbing the normal conditions of its operation. In connection with this, such methods are also applicable for the investigation of living organisms.

On the basis of our analysis of the statistical methods considered, we can formulate the following basic problems related to the estimation of system operators under normal operating conditions:

1) estimating the operators of linear systems with distributed parameters;

2) estimating the operators of linear systems shunted by feedback paths without disturbing normal system operating conditions;

3) estimating the operators of nonlinear systems, in particular, determining the nonlinearities and defining their characteristics;

4) defining the accuracy of the estimates of system operators obtained.

Received November 21, 1958

LITERATURE CITED

[1] R. E. Kalman, Design of Self-Optimizing System, Trans. ASME, vol. 80, No. 2, (1958).

[2] J. H. Laning and R. H. Battin, Random Processes in Automatic Control, New York (1956).

[3] P. E. A. Cowley, The Application of the Analog Computer to the Measurement of Process Dynamics, Trans. ASME, vol. 79, No. 4 (1957).

[4] T. P. Goodman and Reswick, Determination of System Characteristics from Normal Operating Record, Trans. ASME, vol. 78, No. 2 (1956); vol. 79, No. 3 (1957).

[5] H. Thal-Larsen, Correlation Functions and Noise Patterns in Control Analysis, Trans. ASME, vol. 80, No. 2 (1958).

[6] R. A. Woodrow, Trans. Soc. Instr. Tech., (September, 1958).

[7] Yu. P. Leonov and Z. M. Mikhailova, "On an approximate method for synthesizing optimal linear systems for signal separation," [in Russian] Otchet IAT AN SSSR (1958).

[8] Yu. P. Leonov, "On an approximate method of synthesizing optimal linear systems for separating signal from noise," [in Russian], Automation and Remote Control (USSR) 19, 8 (1958).*

[9] I. E. Kazakov, "Approximate probabilistic analysis of operational accuracy of essentially nonlinear systems," [in Russian], Automation and Remote Control (USSR) 17, 5 (1956).*

* See English translation.

A LIST OF THE LITERATURE OF 1958 ON MAGNETIC ELEMENTS FOR AUTOMATION, REMOTE CONTROL AND COMPUTING TECHNOLOGY

1. General Questions

Terminology bibliography, questions of standardization, etc. are in the Materials of the All-Union Conference of the Physics of Magnetic Materials (Leningrad, Dec., 1957). Izv. AN SSSR, ser. fiz. 22, 10 (1958).

B. S. Sotskov, "On the question of the dimensions of electromagnetic elements," Automation and Remote Control (USSR) 19, 9 (1958) pp. 849-854.*

G. V. Subbotina, "List of literature on magnetic and contactless magnetic elements," Automation and Remote Control (USSR) 19, 4 (1958) pp. 379-389.*

R. I. Yanus, "Electrotechnical iron and steel," Fiz. Metllov i Metallovedenie, 6, 2 (1958), pp. 382-384; biblio. 9.

2. Ferromagnetic Materials. Cores

a) Magnetic Materials

O. N. Al'trauzen, "Investigation of the effect of temperature on the magnetic properties of electrotechnical steels," Élektrichestvo 6 (1958) pp. 80-82; 5 figures.

A. A. Glazer and Ya. S. Shur, "On the nature of the effect of thermomagnetic processing of magnetically soft ferromagnets," Izv. AN SSSR, ser. fiz., 22, 10 (1959) pp. 1205-1211; 4 figures, biblio. 9.

G. S. Kandaurova, Ya. S. Shur and E. V. Shol'ts, "The magnetic properties of magnetically anisotropic specimens of ferromagnetic powders. III. Anisotropic magnetic properties," Fiz. Metallov i Metallovedenie, 6, 2 (1958) pp. 229-236; 11 figures, biblio. 8.

E. I. Kondorskii and V. L. Sedov, "Variations in magnetization saturation and electrical impedance of iron-nickel alloys with thorough compression with low temperatures," J. Exptl.-Theoret. Phys. (USSR) 35, 4 (1958) pp. 845-853; 7 figures, biblio. 10, 3 tables.

N. V. Kotel'nikov, "Creation of a magnetic structure by magnetic poles in thin iron films obtained by an electrolytic method," Fiz. Metallov i Metallovedenie, 6, 2 (1958) pp. 222-228; 7 figures, biblio. 9.

V. A. Kutyashev, "On the determination of the magnetic properties of a steel sheet on the entire sheet or on individual layers," Izb. vyssh. uchebn. zaved., Élektromekhanika 4 (1958) pp. 3-12; 9 figures, biblio. 6.

L. V. Mironov, "Structure transformation upon annealing of cold-rolled transformer steel," Izv. AN SSSR, ser. fiz. 22, 10 (1958) pp. 1231-1236; 7 figures, biblio. 5.

M. A. Panasenkov, "Surface effects in massive ferromagnetic bodies with nonlinearities taken into account," Élektrichestvo 9 (1958) pp. 1-7.

I. M. Puzei, "Temperature stabilization of the magnetic properties of alloys," Izv. AN SSSR ser. fiz. 22, 10 (1958) pp. 1194-1199; 8 figures, biblio. 6.

R. I. Rabkin, S. A. Soskin and B. Sh. Epstein, "The synthesis and magnetic properties of ferrites with rectangular hysteresis loops," Izv. AN SSSR ser. fiz. 22, 10 (1958) pp. 1217-1224; 9 figures, biblio. 9.

***See English translation.**

L. I. Rabkin, B. Sh. Épstein and Ya. G. Koblents, "Ferrites with rectangular hysteresis loops," Élektrichestvo 11 (1958) pp. 59-68; 16 figures, biblio. 32.

V. I. Sokolov, "Magnetically soft alloy 50N with high magnetic properties," Elektrichestvo 4 (1958) pp. 70-72; 2 figures, biblio. 6.

Ya. S. Shur, E. V. Shtol'ts and G. S. Kandaurova, "The magnetic properties of magnetically anisotropic specimens of ferromagnetic powders." IV. Temperature dependencies of the magnetic properties of powdered specimens of alloy MnBi, "Fiz. Metallov i Metallovedenie 6, 3 (1958) pp. 420-425; 6 figures, biblio. 9.

b) Magnetization Processes and Dynamic Characteristics of Ferromagnetic Materials (Losses, Eddy Currents, Viscosity, etc.)

L. F. Beits, "On the thermal effects accompanying magnetization processes," Izv. AN SSSR ser. fiz. 21, 8 (1957), pp. 1184-1196.

K. B. Vlasov, "Several questions in the theory of elastic ferromagnetic (magnetostrictive) media," Izv. AN SSSR ser. fiz. 21, 8 (1957) pp. 1140-1148.

P. A. Denesh, "Anisotropic magnetic cores," Izv. AN SSSR ser. fiz. 21, 8 (1957) pp. 1205-1212; 5 figures, biblio. 4.

E. I. Kondorskii and L. G. Smirnova, "Investigation of the frequency dependence of the permeability of certain iron-nickel and iron-cobalt alloys in the frequency range of 10^5 to 10^7 cycles," Fiz. Metallov i Metallovedenie, 6, 2 (1958) pp. 237-246; 9 figures, biblio. 10.

K. M. Polivanov, "Surface effects in ferromagnetic materials with rectangular hysteresis loops with viscosity present," Izv. vyssh. uchebn. zaved., Élektromekhanika, 6 (1958) pp. 3-13; 2 figures, biblio. 6, 1 appendix.

K. M. Polivanov and N. A. Kutyashev, "Surface effects in anisotropic sheets," Iaz. vussh. uchebn. zaved., Élektromekhanika 3, (1958) pp. 3-11; 5 figures, biblio. 1.

M. A. Rozenblat, "Dynamic characteristics of cores with rectangular static hysteresis loops (influence of eddy currents)," Automation and Remote Control (USSR) 19, 1 (1958) pp. 75-84; 7 figures, 3 tables, biblio. 4.*

M. A. Rozenblat, "The dependence of the static characteristics of toroidal cores on their geometric dimensions," Automation and Remote Control (USSR) 19, 8 (1958) pp. 788-799; 11 figures, biblio. 3.*

L. G. Smirnova, "Investigation of the frequency dependence of the permeability of certain iron-nickel and iron-cobalt alloys in the frequency range 10^5-10^7 cycles," Dissertation, TsNII Chermeta (1958) 8 pp.

L. G. Smirnova, "Experimental investigation of the permeability of certain magnetically soft materials at high frequencies," Izv. vyssh. uchebn. zaved., Radiotekhnika (1958) pp. 685-689; 4 figures, biblio. 2.

B. M. Fradkin, "Electrodynamic processes in ferromagnetic bodies characterized by rectangular hysteresis loops," Trudy Mosk. énerg. in-ta, 27 (1958) pp. 179-204; 17 figures, biblio. 6.

c) Construction and Technology of Preparing Cores and Chokes

R. Kh. Bal'yan, "On the design of power transformers with low power at normal frequencies," Vestn. élektroprom-sti, 10 (1958) pp. 41-50.

N. I. Bykov and B. I. Timofeev, "A machine for winding conductors on closed core rings," Class 21g, i_{01}, No. 115375. Byulleten' izobretenii, 10 (1958).

Ya. V. Petrov, "Choosing a rational type of magnetic circuit for frequency triplers," Izv. vyssh. uchebn. zaved., Élektromekhanika 4 (1958) pp. 54-60; 1 figure, biblio. 5.

M. A. Rozenblat and O. A. Sedykh, "Principles of contructing series of toroidal cores for magnetic amplifiers," Standartizatsiya 6 (1958) pp. 3 and 42; 2 figures, 6 tables.

F. F. Sokolov, "On the current density in transformer and choke coils," Tekhn. kino i televideniya, 10 (1958) pp. 52-54.

* See English translation.

A. I. Uzkov and A. M. Byleev, "A machine for winding conductors on closed cores," Class 21 g, 1_{01} No. 110274. Byulleten' izobretenii, 1 (1958).

B. V. Shamrai, "Optical geometric dimensions of toroidal cores," Izv. Lening. Élektrotekh. In-ta, 34 (1958) pp. 149-155.

3. General Questions in the Theory of Nonlinear Magnetic Circuits (Design of Circuits with Steel, Ferroresonance, etc.)

C. P. Andreev, V. E. Boglyubov and V. L. Dyatlov, "On taking into account the actual properties of the materials in Ramey Magnetic amplifiers," Trudy Mosk. Energ. In-ta., 27 (1958) pp. 118-126; 9 figures, biblio. 4.

L. A. Bessonov, "Autooscillation in circuits with nonlinear capacitances," Trudy Vses. Zaochn. Energ. In-ta., 9 (1958) pp. 116-119; 3 figures, biblio 1.

L. A. Bessonov, "The principle of mutuality of feedback flexibility and single-phase autooscillation in systems with magnetic amplifiers," Izv. vyssh. uchebn. zaved., Énergetika 6 (1958) pp. 48-52; 3 figures, biblio. 2.

L. A. Bessonov and M. S. Rybakov, "Magnetizing of ferromagnetic devices by ac at working frequencies, "Trudy Vses. zaochn. Energ. In-ta. 9 (1958) pp. 120-127; 11 figures.

L. A. Bessonov and M. S. Rybakov, "Magnetic devices in asymmetric modes," Trudy Vses. zaochn. Energ. In-ta, 9 (1958) pp. 103-115, 4 figures.

D. I. Bogdanov, "On the theory of a ferroresonant voltage stabilizer," Vestn. Elektropom-sti, 7 (1958) p. 43; 5 figures, biblio. 2.

D. I. Bogdanov and G. K. Evdokimov, "Ferroresonant stabilizers," GEI, Biblioteka po Avtomatike (1958) p. 80; 31 figures, biblio. 12.

B. K. Bul', "Designing magnetic circuits and ac coils from given parameters," Automation and Remote Control (USSR) 18, 10 (1957) pp. 918-926; 2 figures.*

A. A. Grachev, "On the continuous emf spectrum of cyclical magnetic polarity reversal," Izv. vyssh. uchebn. zaved., Radiofizika 1, 2 (1958) pp. 71-78.

Z. Zil'berman, "Conditions of maximum accuracy for ferroresonance voltage stabilizers," Radiotekhnika 12, 2 (1957) pp. 71-72.

P. A. Ionkin, "On designing nonlinear circuits by iterative methods," Élektrichestvo 3 (1958) pp. 45-49; 7 figures.

P. A. Ionkin, "On computing the transient responses in nonlinear circuits by the method of successive ordinates," Izv. vyssh. uchebn. zaved., Elektromekhanika 2 (1958) pp. 43-50; 7 figures, biblio. 8.

P. A. Ionkin, "On designing circuits with mutual inductance and with inertial nonlinear elements," Trudy Mosk. Énerg. In-ta, 27 (1958) pp. 60-66; 4 figures, biblio. 10.

L. F. Kotova, "Use of the method of successive approximations in the design of circuits with nonlinear elements, Trudy Vses. zaochn. énerg. In-ta 9 (1958) pp. 140-145; 5 figures, biblio. 1, 1 appendix.

R. O. Kubo, "On some new achievements in the domain of magnetic investigations in Tokyo," Izv. AN SSSR ser. fiz. 21, 8 (1957) pp. 1197-1204.

B. F. Lastochkin, "Investigation of the conditions for exciting subharmonic oscillations in tuned circuits with nonlinear inductances," Dissertation, Mosk. énerg. In-ta (Moscow, 1958).

B. F. Lastochkin, "Investigation of the conditions for exciting subharmonic oscillations in tuned circuits with nonlinear inductances," Trudy Mosk. Énerg. In-ta. 27 (1958) pp. 205-218; 4 figures, biblio. 11.

A. G. Lur'e, "Theory of ferroresonance voltage stabilizers," Dissertation, Leningrad Polytechnic Institute (1958).

L. D. Mazur, "Designing ferroresonant circuits," Trudy Taganrogsk radiotekhn. In-ta, 3, 2 (1957) pp. 145-154; 2 figures, biblio. 3.

* See English translation.

K. M. Polivanov, "Dynamic characteristics of electrical circuit elements," Dokl. AN SSSR 118, 1 (1958) pp. 80-83; 3 figures, biblio. 4.

V. F. Tabachinskii, "Computing the transient responses in nonlinear circuits by a graphical method of construction in the system of flux of magnetic (electrical) mf (emf)," Sb. Leningr. In-ta, Inzh. zh-d. transp. 161 (1958) pp. 91-101; biblio. 5.

R. I. Yanus and L. A. Fridman, "On cases of incorrect use of magnetization switching curves for approximate design of circuits with ferromagnets," Élektrichestvo 6 (1958) pp. 77-80; biblio. 7.

4. Magnetic Amplifiers. Applications

a) Use in Servo Systems and for Controlling Electrical Machines

D. A. Alenchikov and I. V. Gusev, "Electric drive with an asynchronous motor," Class 21c, 59 No. 115919. Byulleten' izobretenii 11 (1958).

V. D. Afanas'ev and V. A. Ivobotenko, "Speed control of a dc motor by supplying it from a magnetic amplifier," Élektrichestvo 3 (1958) pp. 14-19; 6 figures, 1 table, biblio. 1.

V. D. Baryshnikov, "Differential feedback in automatic control and regulation systems with electromagnetic and magnetic amplifiers," Izv. Leningr. Élektrotekhn. In-ta 34 (1958) pp. 58-70; 12 figures, biblio. 2.

A. V. Basharin, "A graphical method of designing choked magnetic amplifiers for auto-control systems for dc drives," Izv. Leningr. Élektrotekhn. In-ta. 35 (1958) pp. 93-104.

A. V. Basharin, I. N. Vinogradov, A. I. Bychkov and I. B. Byval'kevich, "Motor-generator electric drives with electromagnetic and magnetic amplifiers," Élektrochestvo 11 (1958) pp. 51-54; 8 figures, biblio. 5.

M. V. Belyaev, "Characteristics of motor-generator circuits using magnetic amplifiers," Sb. St. Ural'sk. Politekhn. In-ta. 64 (1958) pp. 147-160.

V. L. Benin and L. N. Senkevich, "Design of a four-choke reversible circuit for an asynchronous short-circuited motor in automatic control systems," Cb. tr. Labor. Gidravl. Mashin. AN SSSR 7 (1958) pp. 198-211.

G. S. Buberman, "On the use of a magnetic drive with an electric motor in the textile and light industries," Avtomatika AN USSR 6 (1958) pp. 21-29; 6 figures, biblio. 6.

I. N. Vinogradov, "Contactless scheme for drive control by a motor-generator eystem," Izv. Vussh. uchebn. zaved. Élektromekhanika 7 (1958) pp. 112-119; 8 figures, biblio. 5.

V. G. Zagorskii, "A device for the automatic control of an asynchronous generator's voltage," Class 21c, 64_{50}, No. 113315. Byulleten' izobretenii, 6 (1958).

F. N. Zuikov and A. I. Khozhainov, "Geometric constructions for an asynchronous motor with a saturable core," Elektrichestvo 11 (1958) pp. 14-18; 6 figures, biblio. 1.

V. I. Klyuchev and V. I. Yakovlev, "The use of magnetic amplifiers for controlling motor-generator systems in excavator electric drives," Élektrichestvo 6 (1958) pp. 59-63; 5 figures, biblio. 1.

E. M. Kovarskii, "Reactive power of a choke controlled electric drive," Vestn. Élektroprom-sti 4 (1958) pp. 21-26; 8 figures, biblio. 3.

I. P. Kopylov, "Connecting a magnetic amplifier and an electrical machine," Élektrichestvo 6 (1958) pp. 56-58; 3 figures, biblio. 8.

I. P. Kopylov, "A dc electric motor, Class $21d^1$, 38 No. 113162. Byulleten' izobretenii 5 (1958).

I. P. Kopylov and V. I. Radin, "Experimental investigation of an electromagnetic amplifier with magnetization," Vestn. élektroprom-sti 2 (1958) pp. 19-23; 7 figures.

V. S. Kulebakin, "A device for starting, and controlling the rotational speed of asynchronous motors and dc motors," Class $21d^2$, 21_{02}, No. 113886. Byulleten,' izobretenii 6 (1958).

V.R. Kulikov, "A method for analyzing and computing the transient responses in the automatic control of generator excitation by means of a magnetic amplifier," Automation and Remote Control (USSR) 19, 6 (1958) pp. 565-573; 7 figures, biblio. 5.

V. R. Kulikov, "A method for analyzing and computing the transient responses in the automatic control of generator excitation by means of a magnetic amplifier," Trudy Khar'kovsk. Polutekhn. In-ta. 17 (1958) pp. 65-75.

A. I. Kupryakova, "A measuring organ for the automatic voltage control of three-phase generators,"Class $21d^2$, 6_{01}, No. 114209. Byulleten' izobretenii 7 (1958).

E. F. Kusov, "Dynamic braking of an asynchronous hoist motor with a saturable choke in the rotor circuit," in book Automation and Electrification in the Coal Industry [in Russian]. Ulgetekhizdat (Moscow, 1958) pp. 101-126.

V.S. Lerner, "A proportional speed block for an automatic control scheme," Vestni. Élektroprom-sti, 10 (1958) pp. 34-36.

V. A. Oleinikov, "Graphical analysis of the transient responses in nonlinear systems when operating in the sampled-data mode," Izv. Leningr. Élektrotekhn. In-ta 34 (1958) pp. 71-78; 8 figures, biblio. 6.

V. A. Oleinikov, "Magnetic amplifier operation in a sampled-data control or regulation system," Izv. Leningr. Élektrotekhn. In-ta. 34 (1958) pp. 46-57; 8 figures, biblio. 3.

N. G. Pereslegin, "Throttle starting of asynchronous motors," Nauchnye doklady vyssei skoly. Gornoe delo 1 (1958) pp. 181-194.

M. A. Rozenblat, "Magnetic amplifiers for automatic control systems," in collection: Automatic Control and Computing Technology, Vol. 1. [in Russian]. Mashgiz (Moscow, no year given) pp. 182-204; 23 figures.

O. B. Rosenbauli and I. N. Selivokhin, "Construction of the characteristics of a dc drive with throttle control," Élektrichestvo 10 (1958) pp. 31-34.

L. L. Rotkop, "A sampled-data method of asynchronous motor speed control by the use of contactless control apparatus," Vestn. Élektroprom-sti 1 (1958) pp. 5-8; 5 figures, biblio. 4.

L. L. Rotkop, "A device for asynchronous motor speed control," Class $21d^2$, 51, No. 111553, Byulleten' izobretenii 3 (1958).

I. A. Sud-Zlochevskii, "A device for voltage and frequency control of a motor-generator transformer," Class 21c, 64_{57} No. 110842. Byulleten' izobretenii 2 (1958).

V. P. Teslenko, "An electronic-ionic system of speed control with magnetic amplifiers," Trudy Khar'kovsk. Politekhn. In-ta. ser. élektromashinostr. 20, 1 (1958) pp. 31-51; 5 figures, biblio. 2.

O. A. Toloknow, M. T. Chirkov and I. A. Erokhin, "A motor-generator system with magnetic amplifiers," Nauchn. dokl. vyssh. shk., mashinostr. i priborostr. 3 (1958) pp. 58-61.

L. F. Shklyarskii, "On the transient responses in an electric drive with throttle control for the hoist of a multi-bucket dredge," Élektrichestvo 11 (1958) pp. 47-50; 7 figures, biblio. 7.

Ya. P. Shtrafun, "A device for the automatic control of a synchronous generator," Class $21d^2$, 6_{01}, No. 113987. Byulleten' izobretenii 6 (1958).

Ya. N. Shtrafun and M. S. Rostovtseva, "Investigation of an automatic excitation system for power turbogenerators with power semiconducting rectifiers," Élektrichestvo 8 (1958) pp. 37-41; 5 figures.

Ya. N. Shtrafun, F. V. Chalyi and E. L. Ettinger, "A device for the self-excitation of synchronous machines," Class $21d^2$, 6_{01}, No. 114208. Byulleten' izobretenii 7 (1958).

b) Use in Current and Voltage Regulators and Stabilizers

A. M. Bamdas, V. A. Somov and V. A. Suchkov, "A welding transformer with smooth voltage control by means of shunt magnetization," Izb. vyssh. uchebn. zeved., Élektromekhanika 8 (1958) pp. 61-65; 4 figures, 1 table, biblio. 4.

S. P. Bykovskii, A. A. Kuritsyn and F. P. Kolomin, "A bearing element for current and voltage stabilizers," Class $21d^2$, $53O_2$, No.1 14949. Byulleten', izobretenii 9 (1958).

M. D. Veinberg, "Ferroresonance voltage stabilizers built of toroidal magnetic circuits," Radiotekhnicheskoe proizvodstvo 15 (1957) pp. 56-58; 5 figures.

I. G. Gol'dreer and Yu. I. Spektor, "An ac voltage stabilizer," Class $21d^2$, 42_{02} No. 115773. Byulleten' izobretenii 10 (1958).

I. G. Gol'dreer and Yu. I. Spektor, "A dc voltage stabilizer," Class $21d^2$, 42_02 No. 115774. Byulleten' izobretenii 10 (1959).

A. G. Ivakhnenko and N. V. Petina, "Voltage stabilizers with combined control," Izdatelstvo AN USSR (Kiev 1958); 246 pp. 114 figures, 60 tables, biblio. 68.

A. A. Moderov, "A current stabilizer with nonlinear electromagnetic elements," Izv. vyssh. uchebn. zaved., Élektromekhanika 7 (1958) pp. 99-111; 10 figures, 1 table, biblio. 4.

G. N. Petrov, S. S. Okun' and B. L. Kraiz, "Smooth contactless voltage control of a loaded transformer," Vestn. Élektroprom-sti 3 (1958) pp. 1-8; 9 figures, biblio. 3.

A. G. Pinchuk, "A voltage stabilizer," Class $21d^2$, 42_{02} No. 114363. Byulleten' izobretenii 7 (1958).

F. F. Sokolov, "Computed power of ac voltage stabilizers," Vestnik Élektroprom-sti 2 (1958) pp. 72-74; 4 figures, biblio. 7.

P. V. Chebyshev, "Ferroresonance voltage stabilizer," Class $21d^2$, 42_{01} No. 113922. Byulleten' izobretenii 6 (1958).

c) Other Applications

I. S. Bruk, "Controllable reactors for long-distance electric transmission," Élektrichestvo 7 (1958) pp. 14-19; 8 figures.

Ya. E.Vol'fkovich, "Regulation of electric furnaces by saturable reactors," Prom. Énergetika 7 (1958) pp. 31-32.

I. V. Gavrilov and V. P. Shipillo, "A device for phase control by ionic transformers," Class $21d^2$, 12_{03} No. 113990. Byulleten' izobretenii (1958).

P. P. Ostryi and V. M. Kvyatkovskii, "A device for controlling an invertor's quenching angle," Class $21d^2$, 12_{03} No. 114006. Byulleten' izobretenii 7 (1958).

M. V. Titarenko, "A device for locating breaks in insulation and determining the degree of grounding in operative dc circuits," Class 21e, 29_{11} No. 112820. Byulleten' izobretenii 5 (1958).

5. Magnetic Amplifiers. Theory, Circuits, Design

a) Books, Monographs, Dissertations

M. A. Boyarchenkov and N. P. Vasil'eva, High-Speed Magnetic Amplifiers. GEI (Moscow, 1958), 32 pp., 53 figures, biblio. 22.

I. S. Gashkovets, "Qualitative theory of operation of three-phase three-choke magnetic amplifier circuits," Dissertation, IAT AN SSSR (1958) (Chekhoslavayskaya AN, Laboratoriya Avtomatiki i Telemekhaniki).

K. E. Dmitriev, "Shaping and designing circuits with magnetic amplifiers," in book, Basis Theory and Maintenance of Analog Computers, chapter II [in Russian]. (Moscow, 1957) pp. 366-416; 17 figures.

V. G. Doroguntsev, "Transformer relays with two electrical input quantities," Dissertation, Mosk. Énerget. In-ta. (Moscow, 1958).

R. A. Lipman, "Questions in the theory and design of high-speed magnetic amplifiers," Dissertation, Mosk. Énerg. In-ta. (Moscow, 1958).

V. I. Molotkov, "Investigation of the operation of a magnetic amplifier with capacitive filters at the output of a rectifier bridge," Dissertation, Leningr. Politekhn. In-ta, (1958).

V. A. Oleinikov, "Analysis of the operation of magnetic and electromagnetic amplifiers and computation of their transient responses in the sampled-data mode of operation," Dissertation, Leningr. Élektrotekhn. In-ta. (1958).

b) Single-Ended Magnetic Amplifiers

R. Kh. Bal'yan and V. I. Laletin, "A device for switching the magnetostriction vibrators of remote hydraulic stands," Class 21a, 48, No. 112120. Byulleten' izobretenii 4 (1958).

A. M. Bamdas and V. A. Somov, "Transformer (autotransformer) with saturable magnetic shunts," Class $21d^2$, 53_{02} No. 110770. Byulleten' izobretenii 2 (1958).

M. V. Belyaev, "A magnetic amplifier's gain and magnetic inertia," Izv. vyssh. uchebn. zaved., Élektromekhanika 6 (1958) pp. 67-76; 6 figures, 1 appendix, biblio. 2.

N. P. Vasil'eva, "Design of power magnetic amplifiers," Lecture for the course "Electromagnetic elements and devices," Vses. zaochn. élektrotekhn. in-ta. (Moscow, 1958) p. 46.

V. S. Vysochanskii, "On varying the load characteristics of magnetic amplifiers," Élektrichestvo 12 (1958) pp. 49-51.

B. B. Gel'perin, "A method for designing saturable chokes," Élektrichestvo 5 (1958) pp. 47-50; 3 figures.

A. D. Drozdov and Yu. I. Karinskii, "A circuit for increasing the multiplicity of the secondary current in magnetic transformer relays," Izv. vyssh. uchebn. zaved. Élektromekhanika 1 (1958) pp. 96-98; 3 figures.

A. G. Zdrok, "On the design of magnetic amplifiers with internal feedback, implemented by transformer steel," Vesti. Élektroprom-sti. 11 (1958) pp. 1-4; 5 figures, biblio. 6.

N. A. Kaluzhnikov, "On the design of a choked magnetic amplifier connected to a single-phase rectifier bridge," Automation and Remote Control (USSR) 19, 3 (1958) pp. 239-256; 15 figures, biblio. 5.*

I. P. Kopylov and V. I. Radin, "A differential electromagnetic amplifier," Class $21a^2$, 18_{01} No. 114299. Byulleten' izobretenii 7 (1958).

B. L. Kraiz, "Choice of parameters and methods of designing transformers with dc magnetization," Izv. vyssh. uchebn. zaved., Élektromekhanika 2 (1958) pp. 69-77.

G. I. Kudryavtseva, "High-speed magnetic amplifiers for servo systems," Élektrichestvo 6 (1958) pp. 41-46; 5 figures, biblio. 6.

G. I. Kudryavtseva, "High-speed magnetic amplifiers," Trudy Mosk. Énerg. Inst-ta. 27 (1958) pp. 243-258; 5 figures, biblio. 6.

V. R. Kulikov, "High-speed magnetic amplifiers for remote control," Élektrichestvo 10 (1958) pp. 38-42; 4 figures, biblio. 4.

R. A. Lipman, "Controlling high-speed magnetic amplifiers by means of semiconducting triodes," Vestn. Élektroprom-sti 1 (1958) pp. 23-31; 9 figures, biblio. 3.

R. A. Lipman and I. B. Negnevitskii, "Choked magnetic amplifiers," Élektrichestvo 6 (1958) pp. 49-55; 7 figures, biblio. 4.

V. I. Molotkov, "Operation of a magnetic amplifier on a rectifier bridge with capacitive filters," Trudy. Leningr. Politekhn. In-ta. 194, Radiofizika (1958) pp. 75-84; 6 figures, biblio. 10.

A. A. Naumov and A. M. Aizen, "Analysis of free-running magnetic amplifiers," Izv. vyssh. uchebn. zaved., Élektromekhanika 7 (1958) pp. 93-83; 2 figures, 2 tables, biblio. 1.

O. B. Rosenbauli, L. É. Spiridovskaya and R. M. Rodin, "Differential transformer magnetic amplifier," Class $21a^2$, 18_{08} No. 115771. Byulleten' izobretenii 10 (1958).

L. L. Samurina, "Characteristics of a single-half-wave magnetic amplifier," Izv. vyssh. uchebn. zaved., Élektromekhanika 9 (1958) pp. 66-72; 5 figures, biblio. 3.

* See English translation.

L. V. Safris, "On the question of the transient responses in magnetic amplifiers with inductive loads connected via rectifiers," Automation and Remote Control (USSR) 19, 3 (1958) pp. 228-238; 18 figures, biblio. 6.*

A. V. Stupin, "On the design of magnetic amplifiers with ac loads," Vestn. Elektroprom-sti 7 (1958) page 47; 6 figures, biblio. 3.

A. V. Stupin, "Special features of designing a magnetic amplifier operating on an inductive load beyond rectifiers, with shunting capacitance present," Vestn. Élektroprom-sti 11 (1958) pp. 5-10; 7 figures, biblio. 3.

N. D. Toroptsev and V. I. Kriventsov, "A device for equal phase loads in three-phase networks," Class $21d^2$, 43 No. 113417. Byulleten' izobretenii 6 (1958)

N. M. Yakovlev, "Computation and determination of temporal parameters of magnetic amplifiers," in Collected Papers on Electric Instrument Construction [in Russian]. Leningradsk. In-ta. Tochnoi Mekhan. i Optiki 28 (Leningrad, 1957) pp. 17-30; 5 figures, biblio. 2.

c. Push-Pull Magnetic Amplifiers

A. N. Gerasimov, "A push-pull magnetic amplifier by a bridge circuit with ac output," Class $21a^2$, 18_{08} No. 113335. Byulleten' izobretenii 6 (1958).

L. V. Maziya and A. A. Stupin, "A push-pull magnetic amplifier," Class $21a^2$, 18_{08} No. 113997. Byulleten' izobretenii 6 (1958).

I. A. Petrusenko, V. G. Baranovskii and V. A. Dunaikin, "A push-pull magnetic amplifier with ac output," Class $21a^2$, 18_{08} No. 111846. Byulleten' izobretenii 2 (1958).

B. V. Shamrai, "Graphical construction of the load characteristics of push-pull magnetic amplifiers with feedback," Izv. Leningr. Élektrotekhn. In-ta. 35 (1958) pp. 105-111.

d. Multistage Amplifiers

O. B. Rosenbauli and R. N. Rodin, "A two-stage magnetic amplifier," Clasd $21a^2$, 18_{08} No. 112022. Byulleten' izobretenii 3 (1958).

E. A. Tanskii, "On trying a magnetic amplifier for a servo system," in Collected Papers on Electric Instrument Construction [in Russian]. Leningradsk. In-ta. Tochno. Mekhan. i Optiki 28 (1957) pp. 64-69; 1 figure.

S. I. Khmel'nik, "A functional transformer of the load current of a two-stage magnetic amplifier circuit," Class $21a^2$, 18_{08} No. 114548. Byulleten' izobretenii 8 (1958).

e) Multiphase Amplifiers

A. M. Bamdas, V. A. Somov and A. O. Shmidt, "Certain variants in the construction of single-phase and three-phase transformers controlled by magnetizable shunts," Izv. vyssh. uchebn. zaved., 10 (1958) pp. 115-123; 15 figures, biblio. 5.

G. G. Demirchyan, "Electromagnetic processes in magnetizable current transformers," Vestn. Élektroprom-sti 10 (1958) pp. 37-41.

M. S. Libkind, "A three-phase controlled reactor," Class $21d^2$, 55 No. 116182. Byulleten' izobretenii 11 (1958).

V. A. Somov, "A three-phase stabilized transformer," Class $21d^2$, 53_{02} No. 111834. Byulleten' izobretenii 3 (1958).

I. M. Stolyarov, "A three-phase magnetic amplifier," Class $21a^2$, 18_{08} No. 113840. Byulleten' izobretenii 6 (1958).

f. AC and High-Frequency Amplifiers

K. S. Volchkov, "Certain optimal relationships in ideal magnetic amplifiers with ac control signals," Automation and Remote Control (USSR) 19, 1 (1958) pp. 85-94; 6 figures, biblio. 2.*

A. L. Mikaélyan, "The problem of creating ferrite amplifiers at super-high frequencies," Radiotekhnika i Élektronika 11 (1958) pp. 1323-1347; 14 figures, biblio. 28.

* See English translation.

V. Yu. Petryn'kin, V. I. Molotkov, L. N. Pakhomov and S. G. Pellikan, "A low-power audio-frequency magnetic amplifier of ferrites," Trudy Leningr. Politekhn. In-ta, Radiofizika pp. 69-74; 7 figures, biblio. 3.

6. Magnetic Voltage Amplifiers, Modulators, Sondes (Magnetometers)

G. I. Devidze and I. M. Chizhenko, "A device for measuring large dc currents," Class 21e, 36_{01} No. 113957. Byulleten' izobretenii 6 (1958).

B. Ya. Zhukhovitskii, "Analysis of choke operation in modulator circuits," Izv. vyssj. uchebv. zaved., Radiotekhnika 4 (1958) pp. 475-479; 8 figures, biblio. 1.

F. I. Kerbnikov and M. A. Rozenblat, "Magnetic modulators with lateral excitation," Automation and Remote Control (USSR) 19, 9 (1958) pp. 836-848; 9 figures, biblio. 11.

Yu. I. Spektor, "On the transformation characteristics of magnetic amplifiers (sondes) of the second harmonic type in the free-running mode of operation," Avtomaticheskii Control' i Izmeritel'naya Tekhnika 2 (Kiev, 1958) AN USSR.

R. Sh. Shklyar, A. A. Popov and A. L. Salov, "A new magnetometer with ferromagnetic sondes," Zavodsk. Laboratoriya 24, 7 (1958) pp. 892-893.*

D. A. Shturkin, "Comparative testing of ferrosondes with longitudinal and with lateral excitation" Trudy In-ta. Fiz. Metallov 20 (1958) pp. 111-124.

7. Magnetic Digital Elements

a) Books, Dissertations

Z. I. Zavolokina, Magnetic Elements in Digital Computers. GEI (Moscow, 1958) 119 pp. 67 figures, biblio. 46.

E. M. Martynov, Contactless Switching Devices. GEI (1958) 80 pp., 43 figures, 2 tables.

I. P. Pritaka, "Investigation and Development of remote protective relays for agricultural electric stands," Dissertation, Academy of Agricultural Science, USSR (Kiev, 1958).

M. N. Sinyagina, The Application of Remote-Control Contactless Switching Devices for the Objectives of Communal Farms. Author's Report, Cand. Diss. K. D. Panfilov, Acad. of Communal Farming, p. 13 (1958).

N. M. Sotina and Ya. A. Khetagurov. The Basic Elements and Some Circuits of Computing Devices, Izdatelstvo ITM i VT AN SSSR (1958) 25 pp.

b) Surveys and Applications of Digital Elements in Control Systems and Computers

A. Ya. Artyukhin and V. Z. Khanin, "A single-cycle magnetic shift register," Automation and Remote Control (USSR) 19, 10 (1958) pp. 977-987; 15 figures, biblio. 9.*

A. V. Karibskii, "Principles of constructing magnetic memory devices (a survey)," Sbornki Avtomatiki i Telemekhaniki. Izdatelstveo AN SSSR (1958) pp. 34-40; 3 figures, biblio. 10.

V. V. Kobelev and Yu. I. Vizun, "Program-controlled magnetic memory devices," Trudy ITM i VT AN SSSR (1958) 18 pp. 14 figures.

V. A. Leokene, "A method of constructing a push-pull magnetic chain circuit," Class $21a^4$, 74 No. 111873. Byulleten' izobretenii 3 (1958).

V. A. Leokene, "A method of constructing a push-pull magnetic chain circuit," Class $21a^4$, 74, No. 111874. Byulleten' izobretenii 3 (1958).

V. P. Sabadashev and V. G. Gerasimov, "A contactless magnetic device for the automatic and semiautomatic synchronization of a synchronous generator," Izv. vyssh. uchebn. zavod., Élektromekhanika 10 (1958) pp. 104-114; 13 figures, biblio. 6.

I. M. Shenbrot, "Ferrite-transistor cells in control circuits," Izv. AN SSSR, Otd. tekhn. n. 10 (1958) pp. 18-26; 8 figures, biblio. 7.

* See English translation.

c) Elements in the Form of Magnetic Amplifiers with Feedback

O. I. Aven and S. M. Domanitskii, "A contactless reversible executive mechanism," Class 21c, 59_{34} No. 115772. Byulleten' izobretenii 10 (1958).

Si-shi Van', "The use of contactless elements in linear circuits for dispatcher centralization." Dissertation, Leningr. In-ta. Zh-D Transp. (1958), 12 pp.

L. I. Gutenmakher and I. N. Popov, "A device for protecting electrical equipment (stands)," Class 31c, 68_{50} No. 113119. Byulleten' izobretenii 5 (1958).

D. Ya. Dusavitskii, "A contactless magnetic relay," Class 21g, 4_{05} No. 113127. Byulleten' izobretenii 5 (1958).

V. A. Zhozhikashvili and K. G. Mityushkin, "Relay phenomena in ring circuits containing magnetic cores with rectangular hysteresis loops," Automation and Remote Control (USSR) 19, 1 (1958) pp. 64-74; 10 figures, biblio. 1.

E. P. Krasovskii, V. V. Lauėr and A. G. Shapovalenko, "An ac time relay using magnetic amplifiers," Izv. vyxxh. uchebn. zaved. Énergetika 11 (1958) pp. 59-64; 4 figures, biblio. 2.

I. Kovachik and V. E. Polyakov, "Contactless remote-disconnect apparatus for relay protection circuits in automated energy systems," Izv. vyssh. uchebn. zaved., Énergetika 11 (1958) pp. 15-20.

R. A. Lipman and I. B. Negnevitskii, "Transient responses in magnetic amplifiers operating in the relay mode," Élektrichestvo 7 (1958) pp. 25-30; 5 figures, biblio. 5.

O. A. Maevskii, "A relay device," Class 21 g, 4_{01} No. 113672. Byulleten' izobretenii 6 (1958).

V. I. Novash and L. L. Chervinskii, "A "multicontact" magnetic relay with a choke-transformer magnetic amplifier," Izv. vyssh. uchebn. zaved., Energetika 12 (1958), pp. 10-16; 7 figures, biblio. 4.

I. N. Popov, "Sensitive shielding of a generator stator," Class 21c 68_{50} No. 113283. Byulleten' izobretenii 5 (1958).

I. V. Prangishvili, "A contactless magnetic relay," Class 21g, 4_{05} No. 111529. Byulleten' izobretenii 3 (1958).

I. P. Pritaka, "Graphic Interpretation of Contactless Remote Time-Delay Relays." Avtomatika, AN UkSSR, 1, 94-100, 6 figures, biblio. 6.

M. A. Rozenblat and S. A. Dokhman, "A contactless magnetic relay for the automatic inspection and sorting of manufactured articles," Élektrichestvo 1 (1958) pp. 45-48; 3 figures, biblio. 4.

O. B. Rosenbauli and R. N. Rodin, "A protective device for three-phase electrical stands," Class 21c, 68 No. 113116. Byulleten' izobretenii 5 (1958).

N. I. Chicherin, V. G. Leskov and V. K. Aretm'ev, "A contactless three-position relay," Class 21 g, 4_{05} No. 111155. Byulleten' izobretenii 3 (1958).

d) Hysteretic Elements

G. V. Burdenkov, "A telemetry system with pulse-code modulation," Automation and Remote Control (USSR) 19, 1 (1958), pp. 55-63; 11 figures, biblio. 5.

V. A. Zhozhikashvili and K. G. Mityushkin, "Contactless devices for telecontrol and telesignalling made of magnetic elements," Trudy VNIIE 7 (1958) pp. 56-77; 12 figures, biblio. 2.

V. A. Zhoshikashvili and K. G. Mityushkin, "A single-cycle pulse distributor made of elements with rectangular hysteresis loops," Trudy VNIIE 7 (1958) pp. 131-149; 11 figures, biblio. 2.

I. V. Prangishvili, "A contactless change-over switch," Class $21a^3$, 17_{10} No. Byulleten' izobretenii 10 (1958).

G. D. Smirnov, Electronic Digital Computers, GEI (1958) 88 pp., 34 figures.

V. N. Tutevich, "A remote control system built of magnetic elements with rectangular hysteresis loops," Élektrichestvo 11 (1958) pp. 72-73; 1 figure, biblio. 3.

e) Ferroresonance Elements

Ya. G. Koblentx and D. A. Yakovenko, "A contactless magnetic trigger," Class 21g, 4 No. 111358. Byulleten' izobretenii 2 (1958).

8. Magnetic Frequency Generators and Transformers

N. A. Galochkin, "The basic analytical design of a frequency tripler for an active-capacitive load," Sb. Trudy Ivanovsk. Énergetich. In-ta 8 (1958) pp. 278-294; 6 figures, biblio. 2.

A. V. Kalyaev, "Graphicoanalytic calculation of certain modes of frequency-doubler operation," Trudy. Taganrogsk. Radiotekhnich. In-ta 2 (1958) pp. 125-135; 5 figures, biblio. 5.

N. G. Maksimovich and Yu. E. Batranin, "A method for the analysis and design of multiphase ferromagnetic frequency multipliers," Izv. vyssh. uchebn. zaved., Énergetika 10 (1958) pp. 49-58; 6 figures, 1 appendix, biblio. 2.

K. G. Mityushkin, "A magnetic frequency divider," Trudy VNIIE 7 (1958) pp. 157-172; 12 figures, biblio. 1.

A. M. Pshenichnikov, "A device for transforming dc voltage into ac pulses whose frequency is proportional to the original dc," Class 21a, 1 No. 110309. Byulleten' izobretenii 1 (1958).

9. Magnetic Measurements and Methods of Testing Magnetic Elements

V. V. Abramov, "A device for automatic selection of cores by their magnetic properties," Class 213, 37_{67} No. 114214. Byulleten' izobretenii 7 (1958).

A. V. Evgrafov, "A machine for determining the magnetic characteristics of ferromagnetic metals by a ballistic method," Class 21e, 12 No. 115706. Byulleten' izobretenii 10 (1958).

K. É. Zaidel' and G. I. Sharokhin, "Separate measurement of odd and even harmonics of nonsinusoidal curves," Trudy Mosk. Énerget. In-ta. 25 (1958) pp. 127-132; 4 figures, biblio. 5.

M. V. Ivanov, "A method for measuring magnetic inductance or magnetic field strength," Class 21e, 12 No. 113117. Byulleten' izobretenii 5 (1958);

V. A. Kadkin, "A device for measuring the magnetic properties of ferromagnetic materials," Class 213, 12, No. 110228.

I. I. Kifer and B. N. Kozintsev, "A device for demagnetizing rings," Class 21g, 5 No. 114293. Byulleten' izobretenii 7 (1958).

I. I. Kifer, B. P. Kozintsev and E. V. Nitusov, "A device for demagnetizing rings," Class 21g, 5 No. 114307. Byulleten' izobretenii 7 (1958).

L. S. Levin and E. N. Chechurina, "Device UFM-1 for testing magnetic materials," Izmerit. Tekhnika 4 (1958) pp. 64-68.

V. I. Molotkov, "A method for determining the dynamic hysteresis loop by means of two voltmeters," Trudy Leningr. Politekhn. In-ta. 194 (1958) pp. 85-87; 3 figures, biblio. 5.

V. I. Molotkov, "A device for determining the limiting dynamic hysteresis loop of ferromagnetic materials," Class 21e, 12 No. 112088. Byulleten' izobretenii 3 (1958).

V. I. Perov, "A device for measuring electrical hysteresis loops in magnetic materials," Class 21e, 12 No. 116052. Byulleten' izobretenii 11 (1958).

M. A. Rozenblat, "A field-stand for determining the magnetic properties of wire and tape sound conductors," Trudy Vses. N-I. In-ta. Zvukozapisi (1957) pp. 122-128; 4 figures, 2 tables.

N. M. Rudnyi, A. Z. Veksler and A. I. Bulanova, "A method hysteresis and eddy current losses with twofold magnetization," Referaty Nauchno-Issledovatel-skikh Rabot Institutov, Komitet Standartov, Mer i Izmeritel'nykh Priborov pri Sovete Ministrov SSSR, VNIIM 2 (1958) pp. 104-105.

Compiled by G. V. Subbotina and I. S. Trefilova